Three Essays on Demand Estimation

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

András Imre Péchy
from Hungary

approved in July 2016 at the request of
Prof. Dr. Michelle Sovinsky
Prof. Dr. Gregory Crawford

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 20.07.2016

Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgements

The four years spent working on my PhD thesis at the University of Zurich have been a great journey. I had the chance to work with fantastic people, whom I would like to thank here.

I am extremely grateful to my advisor Prof. Michelle Sovinsky. By frequently challenging my ideas, she encouraged me to understand problems in detail and search for the best possible solutions. Her sound advice regarding the art of writing and presenting have been very valuable.

I am indebted to Prof. Gregory Crawford. I benefited from his presence in our Department in terms of both the knowledge and the energy that he has brought along. In particular, his detailed feedbacks were very much appreciated.

I would also like to thank my co-authors Hwa Ryung Lee and Prof. Alon Eizenberg. Working with them was very enlightening, whether it was creative brainstorming or developing the econometric approaches of our projects.

When blocked on a specific question, I often turned to my colleagues Chloé Michel, Lachlan Deer, Nicola Pavanini and Steve Heinke who provided great help in viewing problems from a different angle, and to move forward.

During my undergraduate studies, some people have stimulated my curiosity for studying Economics. Without them, this thesis would not have been written. I am grateful to Prof. Thomas Von Ungern Sternberg, Prof. Rafael Lalive and Prof. Tomaso Duso.

Finally, I would like to thank my parents for their continuous support and encouragement since 1984.

András Péchy, Zurich, February 2016

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Dissertation Overview

Demand estimation has been a major theme in economic research since the beginning of the twentieth century. For example, in 1927, E.J. Working was already discussing the identification of parameters of the demand function[1]. Over the years, publications relating to demand estimation have come to cover vast areas in terms of the economic question, econometric methodology, empirical applications or policy implications. In order to position the contributions of this thesis, I divide the existing literature into two groups. I define the first group to encompass those contributions which aimed at improving demand estimation itself, by answering the question "What are the drivers of the choices economic agents are observed to make?". I define the second group to include those contributions which built on existing demand estimation models and have used them as tools to uncover other economic behaviours of interest. The three independent chapters which compose this thesis can be classified into these two groups as follows. Chapter two belongs to the first group: I propose a method to improve the quality of demand estimation on market level data by augmenting the model with micro data purchase histories. Chapter three (co-authored with Alon Eizenberg and Michelle Sovinsky) can be classified in the second group:

[1] E.J. Working (1927) "What Do Statistical Demand Curves Show?", *The Quarterly Journal of Economics*, 41(2), 212-235.

we use a reduced form demand estimation to empirically assess the effects of vertical restraints on market outcomes. Chapter four (co-authored with Hwa Ryung Lee and Michelle Sovinsky) also belongs to the second group: we build on a structural demand estimation to detect non-price predatory behaviour. Finally, chapter five contains my curriculum vitae. The following subsections provide a summary of each chapter.

1.1.1 Summary of Chapter 2

"Improving Market Level Demand Function Fit using Micro Level Purchase Histories "

In this chapter, I propose a method to augment a demand function estimated on market level data with micro data purchase histories in order to improve the demand function fit. The model builds on the micro data to pull information on loyal and non-loyal households. A loyal household is defined to have a marginal utility of consumption if the good chosen today is the same as the choice of last period. For a non-loyal household, the marginal utility of consumption will not be affected by last period's choice. In a first step, a micro model is defined at the household level, where the dataset is available for a subset of households only. This allows the identification of state dependence, thereby distinguishing loyal and non-loyal households. In the second step, a market level demand function which accounts for state dependence is estimated on all markets. It is based on the usual two moments of the Berry, Levinshon and Pakes (1995) model: predicted market shares must match observed ones and predicted unobserved product characteristics must be orthogonal to instruments. To allow this model to account for state dependence, three modifications are implemented. The state dependence coefficients are taken from the first step estimation and treated as data. A proxy for the state variable (which is not defined in a market level model) is introduced, based on forward iteration of predicted choice probabilities. Finally, an extra moment condition is added, which ensures that the share of loyal transitions predicted by the market level model matches the one observed in the micro data. This second step is estimated in a GMM framework. The resulting market level demand function is based on all market level information available and is actually accounting for the loyal behaviour of consumers observed in the micro data. I show on Monte Carlo simulated datasets that the model improves market share fit. I apply the model to data on yoghurt purchases in the US. I find that the model this empirical application does not perform as well as in the Monte Carlo simulations. I suggest that this mis-performance is likely an indicator of households having very different loyal/non-loyal behaviours in the markets under study.

1.1.2 Summary of Chapter 3

"Technology Adoption, Vertical Restraints and Partial Foreclosure: Changing the Structure of an Industry "

Joint work with Alon Eizenberg and Michelle Sovinsky

Vertical restraints between upstream suppliers and downstream customers often raise anti-competitive concerns. While the extant empirical and theoretical literature focuses on static analyses of such arrangements, this chapter empirically documents their effect in a dynamic environment. Our study focuses on the x86 processor industry during the years 2002-2010 which saw an incumbent, dominant upstream supplier (Intel) attempting to maintain its dominant position versus a smaller contender, Advanced Micro Devices (AMD). Intel's strategy included a controversial program, "Intel Inside", through which it offered its downstream clients rebates and subsidies that were conditioned on the volume purchased from it and, sometimes, on the volume purchased from AMD. This prompted substantial legal action by AMD and competition authorities that finally compelled Intel to curb the scope and nature of the program. We use this observed variation to study the effect of the program on the downstream adoption of AMD's technology. Our analysis integrates several datasets, including detailed sales and prices of PC products, rich details about the provisions of the heterogeneous vertical restraints imposed by Intel on various downstream clients, a comprehensive account of the lawsuits and complaints launched against Intel, and data on the evolution of technology and capacity. These data allow us to investigate the impact of the Intel Inside program using both linear and nonlinear dynamic panel regression methods, delivering similar conclusions.

Our results indicate that the adoption of AMD's technology by a given downstream customer responds negatively to (i) the extent of Intel Inside payments to the downstream customer itself,

(ii) specific restrictions on the extent of usage of AMD's technology imposed by the arrangement, and (iii) a measure of Intel's technology. We further find that the extent of adoption responds positively to (i) a measure of AMD's technology (ii) a measure of AMD's capacity, and (iii) the extent of "anti-Intel Inside" litigation. These findings reflect the importance of dynamics in the technology adoption process: downstream customers weigh the potential benefits from adopting AMD's technology against its costs, reflected in losing benefits from the Intel Inside program. Evaluation of both the benefits and the costs crucially depends on firm's expectations regarding: AMD's production capacity and technological progress contrasted with that of Intel and the future viability of Intel's vertical restraints. The empirical evidence is therefore consistent with a nuanced form of foreclosure: one targeted not at eliminating a rival, but rather at keeping the rival below a threshold of production. In innovative industries that require substantial investments, this may successfully block the rival from developing the capacity that would allow it to threaten the incumbent's dominant position.

1.1.3 Summary of Chapter 4

"Assessing an Efficiency Defense: The Case of Intel's Marketing Campaign "

Joint work with Hwa Ryung Lee and Michelle Sovinsky

Antitrust authorities typically try to establish exclusivity and the anticompetitiveness of loyalty rebates through pricing, but do not address the strategic use of advertising and, more generally, marketing campaigns. In this chapter we focus on non-price anticompetitive behaviour arising from marketing. We propose a Test of Advertising Predation (TAP) that can be used to detect non-price predatory behaviour. The TAP test is based on a structural approach and allows us to disentangle the potential positive impact of a marketing program from the anticompetitive predatory effect. We apply the TAP test to the Intel case, but it can be used to guide antitrust authorities in future cases, as it provides a more general framework for testing for the anticompetitive use of marketing campaigns. We use the test to examine whether Intel's choice of processor marketing via PC firms is consistent with predatory behaviour, and find evidence that the "Intel Inside" marketing campaign had predatory effects.

Chapter 2

Improving Market Level Demand Function Fit using Micro Level Purchase Histories

Acknowledgements: I would like to thank: Michelle Sovinsky, Gregory Crawford, Alon Eizenberg, Brett Gordon, Marc Rysman, Federico Ciliberto, Nicola Pavanini, Chloe Michel, Lachlan Deer, Sergio Maffioletti, Jeremy Fox and Gunter Hitsch. Moreover, I would like to thank participants at internal seminars at the University of Zurich, DIW Berlin and Zurich Workshop in Economics 2015.

2.1 Introduction

In various industries, consumers make regular purchases over time from the same set of goods. Thus, consumers' purchases are linked dynamically and in many cases they are observed switching from one brand to another. With market level data, only those switches that lead to market share changes can be accounted for. However, market level data does not allow us to account loyal/non-loyal behaviour of consumers.

Let us take a very simple example: consider a world with only two goods A and B. Say that the market level data show that market shares of each product are 50% in period 1, and 50% in period 2. Therefore one might conclude that the market is fairly stable, consumers seem to be loyal to the products. However, assume that micro data at the consumer level is available, which reveal that 10% of consumers have switched from A to B and 10% have switched from B to A. This simple example illustrates that market level data may be misleading, as 20% of consumers are actually not loyal.

The goal of this paper is to propose a solution for this issue, by showing how to incorporate loyalty while not observing micro data on all markets. The model is based on a demand function estimated at the market level while accounting for the loyal/non-loyal behaviour of consumers observed in the micro data (household purchase histories) available for a subset of markets. As it will be discussed below, this change can improve the fit of the market level demand function in situations where loyal/non-loyal behaviours are frequent.

For this new information on household purchase histories to be valuable, it has to be accounted for in the demand function. I follow the literature on household level data (see for example Erdem (1996))by adding a state dependence term to the utility function. To reproduce the different purchase patterns observed in the data, I allow for two types of consumers with respect to this state dependence term. For the first type, the loyal consumers, the marginal utility of consumption will be larger if the good chosen today is the same as the choice of last period[1].

[1] Heckman (1981) emphasizes that, if unobserved consumer heterogeneity is not correctly accounted for, the

For the second type, the non-loyal consumers, the marginal utility of consumption will not be affected by last period's choice. This paper proposes a model which, although based on a market level approach (closely following Berry, Levinsohn and Pakes (1995)(hereafter BLP)) provides parameter estimates (and in particular price elasticity estimates) which are net of the effect of state dependence.

The inclusion of the state dependence term in the utility function is important as omitting it can significantly affect the parameter estimates, and thus the fit of the demand function. Suppose the following data generating process: the consumer has purchased good A because this good maximizes his utility which is the sum of the utility stemming from the price, the product characteristics and of the pure state dependence term (assume the consumer is of loyal type, and he purchased good A in the previous period). Suppose you estimate your demand function on this data with a standard BLP, hence the last term is not modelled. Which parameters will be likely to pick up its effect? The consumer heterogeneity coefficients of product characteristics are the most likely to be affected by the omission of this variable. This is because the agent is already more likely to choose product A as his tastes match product A's characteristics. Thus, in the model, a shift in the heterogeneous taste parameters would increase the probability of purchase product A for the types of consumers who were already consuming it. The unobserved product heterogeneity is another candidate to pick the effect up, as it covers everything that the explanatory variables could not account for. However, in this case, it would imply a shift in the utility of consumption for all consumers, not just those who were choosing product A the period before. Therefore, the effect is less likely to be picked up by the unobserved product heterogeneity. Finally, the effect of the state dependence term could be picked up by an autocorrelation of the error terms for a given good and given consumer type. Therefore, the assumption of identically and independently distributed error terms needed for the estimation of the BLP model would not hold in that case.

The model introduced in this paper is useful to practitioners for the following three reasons. First, because managers and competition authorities alike need precise demand function estimates

state dependence term will likely pick up some of its effect. In this paper, I try to account as well as possible for the unobserved consumer heterogeneity, using discrete distributions of unobserved consumer types.

to evaluate various policies of interest, for example the chances of the entry of a competitor into the market. This event is depending on the loyalty of consumers (since a larger share of loyal consumers implies a more difficult product launch for the entrant) thus the demand function modelled should account for this behaviour. Second, the model can be applied to various industries (retail, pharma,..) as the data needs are not too stringent (market level data on all markets, micro data on a subset of markets). Finally, the modelling of state dependence is particularly useful to managers. They can exploit the knowledge about the share and intensity of consumer loyalty to design inter-temporal price or advertisement strategies[2].

The model is estimated in two steps, which relate to distinct datasets: the market level dataset covers all markets while the micro level data is available for a subset of these markets. In the first step, the micro data set is used to estimate the state dependence coefficients of the households exploiting the observed switching patterns. The second step specifies a market level demand function on all existing markets and incorporates these estimated state dependence coefficients into the utility function. To achieve this, state variables are proxied via forward simulation for each consumer type. Moreover, to ensure that the market level model mimics switching probabilities in the right way, the share of loyal purchases observed in the micro data must match by the predictions of the model. The second step is estimated in a General Method of Moments framework (hereafter GMM)[3].

In the first step, a demand function is estimated on the micro data set, a panel data at the household level. The utility function is specified to take advantage of the fact that individuals' consumption choices are observed over a long period of time and thus accounts for *consumer inertia* via two different channels: consumer heterogeneity and state dependence. Consumer heterogeneity is accounted for by modelling several sets of taste parameters, similar to the approach in Berry, Carnall and Spiller (1996). For example for the price coefficient, three types of households can be modelled: low, mid and high price sensitive. State dependence is accounted for with an indicator function evaluating whether the consumption choice of today is identical to the

[2] For a discussion of these implications, see for example Dube, Hitsch and Rossi (2010).
[3] Table 2.A.1 in the appendix provides an overview of the two steps of the model.

choice of the previous period[4]. The coefficient on the state dependence term is also allowed to be heterogeneous, in order to model the two household types: loyal and non-loyal. For the loyal households, the state dependence effect will be positive. While for the non-loyal ones, the effect is normalized to zero. The household is assumed to be myopic, it does not consider the effect of its choice today on the utility it can receive in the next period[5]. Then, for each household, the probability of choosing a product conditional on its state variable can be computed. These probabilities are then estimated using the Maximum Likelihood methodology. The parameters of the state dependence are retrieved and considered as data in the second step of the model.

In the second step, a discrete choice model is estimated at the market level, on the complete set of markets. The utility function is defined similarly to the one of the first part, except for two elements. The state dependence coefficients obtained from the first step are used and treated as data[6]. The state dependence indicator function can not be built based on the market level data, since state variables are not defined in a market level framework. Therefore, the state variable is proxied by using its distribution, i.e. the choice probabilities predicted from the previous period (following Eizenberg and Salvo (2014)). This allows me to predict market shares, which now incorporate the state dependence effect. The model is estimated based on BLP's two moment conditions: the predicted market shares must match the observed ones and the predicted unobserved product characteristics must be orthogonal to the instruments.

The goal is to generate a market level demand function that accurately reproduces the switching patterns observed in the micro data, but so far the first and second steps of the model are only linked by the taste parameters. Following Petrin (2002), I add an extra moment condition to the model: that the share of loyal purchases generated by the market level demand function

[4] Other definitions of loyalty have been proposed in the literature. See for example, Guadagni and Little (1983) or Setharaman (2004)

[5] While the forward looking behaviour of households has been shown in various papers (for example Hendel and Nevo (2006)), this assumption is not taken here since the benefits of the model presented in this paper can already be shown by assuming myopic households. Adapting the model to account for forward looking behaviour is an interesting path for future research.

[6] As it is well known in the literature (for a discussion, see Newey and McFadden (1994)), treating parameter estimates as data in the second step of a two step estimation affects the standard errors of the parameters remaining to be estimated. To my knowledge, there is no short-hand solution for the steps used in this paper (first step maximum likelihood, second step GMM) therefore the correction of the standard errors is left for future research.

match those observed in the micro dataset. The model is estimated in a GMM framework.

I analyse the performance of the model in two different environments: once on Monte Carlo simulated datasets and once on products in the fast moving consumer goods industry. This industry is well suited for the model described here because consumers make regular repeated purchases on a stable set of products, and precise estimation of demand functions is a crucial issue. Consumer inertia is a very typical pattern observed in households' purchase histories (see for example Dube, Hitsch and Rossi (2010)). Furthermore, the data in this industry consist of market level data on many markets (for example cities) and household level data. The latter is more difficult to get, but usually available for a few of these markets in sufficient sample sizes. I use Nielsen RetailScan and Nielsen HomeScan datasets[7]. I select the yoghurt product category because it is a product that is quite frequently purchased and offers a large variety of brands. I estimate the model on two Midwest cities in the US.

The estimation results are compared to the BLP model as a benchmark. The results over the Monte Carlo simulated data confirm that augmenting a market level demand function with loyalty parameters estimated on another market does improve the fit of the data. The improvement measured in Mean Squared Error over market shares is estimated between 7.5% to 2%, depending on the setup of the simulated datasets. The difference is driven by the data generating processes: the more different the households populating the two markets (with/without household data) are, the smaller will be the advantage over the benchmark BLP. The results in the empirical application confirm these findings: the model performs 8% worse than the benchmark BLP.

The coefficients of price and product characteristics remain almost unaffected by the addition of the loyalty parameters. This model requires the assumption that the driving force of households' loyal/non-loyal behaviour has to be similar for the market with household data as well as for the market to which the estimates are taken along. However, the setup for which the model is designed excludes the availability of household data on the market of interest. Therefore, if the model fails

[7] Calculated (or derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Information about the data and access are available at http://research.chicagobooth.edu/nielsen/.

to perform better than the benchmark BLP, it might indicate that the market from which the loyalty parameters are taken along is populated with households having too great a difference in characteristics from those of the market of interest. Finally, I present a possible way to proceed in case of this negative outcome. For the case where household data is available for multiple markets, I suggest a simple application of the 2-step model to identify the loyalty coefficient estimates which best perform on the market of interest.

Several papers in the existing literature have aimed at improving market level demand functions to account for switching patterns and state dependence without using micro data. Horsky, Padvlidis and Song (2012) propose a simple strategy: they rewrite conditional purchase probabilities by using Bayes rule as a function of unconditional probabilities which allows them to proxy the probability of a purchase as a function of the observed market shares. They note that their methodology is not well suited to account for consumer heterogeneity, and thus do not claim to identify true state dependence. They apply their model to the salty snack category and find that accounting for the state dependence term leads to more elastic price and affects the estimates of coefficients of advertising variables. Another approach taken is to simulate the transition probabilities of consumer choices. Chen and Yang (2007) and Musalem, Bradlow and Raju (2009) both propose a Bayesian econometrics framework (based on two different data augmenting techniques) to simulate the unobserved purchase probability of the simulated consumers. This approach crucially avoids the integration over all the set of possible consumption histories for all consumers. The model is then estimated in a Markov Chain Monte Carlo (MCMC) method. Switching patterns are simulated until the observed market shares are matched. They claim that tehir model allows for the analysis of consumer level behaviour when only market level data is available. In a frequentist framework, Shcherbakov (2009) develops a dynamic model to estimate switching costs in the cable TV industry. The consumers are simulated based on the observed market shares, and the optimal choice of TV provider is made accounting for future utility levels, and individual heterogeneity. These approaches are of course interesting in cases where data at the micro level is not available. But when micro data is partially available, one can exploit it to improve the fit of the demand function.

Another strand of the literature has focused on using micro level data to identify pure state dependence. Sudhir and Yang (2014) propose an innovative identification strategy for to disentangle state dependence and unobserved heterogeneity. Their model requires a setup where the consumer's choice, as well as the product that is consumed (the two not necessarily matching) are both observed. This is the case in their empirical application to car rentals where free upgrades are a frequent practice. This enables them to observe exogenous shocks to the state variable and identify true state dependence. Dube, Hitsch and Rossi (2010) estimate a demand function accounting for state dependence over a panel data of households, for one large Midwestern city. In order to increase the chances of their model identifying state dependence, they use a very flexible parametrization of individual heterogeneity: a 5 components mixture of normal models. The authors estimate the model in a Bayesian framework using an MCMC algorithm. They find that even with such a flexible representation of heterogeneity, the state dependence they estimate remains robust. Finally, they provide several tests that suggest that the estimated state dependence is due to loyalty, and not to learning or search costs. Their analysis provides a parsimonious demand function for the specific market where household purchase histories are available. However, it is not clear whether the predictions of this model can be used on other markets where household data are not available. A problem to which the model presented in this paper proposes a solution.

Models using both market and micro level data to account for state dependence when estimating demand functions are scarce. To my knowledge, solely Eizenberg and Salvo (2014) have made such an attempt, in the case where micro data is available in the first observed period exclusively. In their paper, they shed light on how the fast changing socio-economic structure of Brazil has affected the landscape of the soda market. They develop a demand model on market level panel data where state dependence is defined at the brand type level (premium, generic or no soda consumption). The authors represent heterogeneity of consumers using discrete types which are defined on the socio-economic groups and past consumption habit. The micro data allows the estimation of the distribution of types in the population for the first period. It is used as the distribution of state variable for the second period and thus the market shares of the second period can be estimated. Then, these type-specific market shares are used to predict the distribution of state variable for the third period. Because micro data is not available for further periods, the authors compute the prediction of the state variable by using data on socio-economic evolution of the population and assumptions on how agents change their consumption behaviour while moving up/down the social ladder. Thus, they obtain a demand function that accounts for state dependence and can be estimated over all periods of the market level data.

The current paper differs from Eizenberg and Salvo (2014) both in terms of the purpose and the approach. In their case, since household panel data is not available for this emerging country, their model relies on strong aggregate variation in the socio-economic characteristics of the population, beside variation in prices, to identify the state dependence coefficient. In particular, this means that state variables are not *observed*, but *predicted*. The current paper is developed for the case when household panel data is available for a limited number of markets. Crucially, the identification of the state dependence coefficients in the first step is based on the *observed* state variables of households. Furthermore, the micromoment ensures in the second step of the model that predictions for the share of loyal purchases are close to the observed ones. The model described in this paper should therefore bring more robust estimates for the state dependence coefficients, and thus improve the fit of the demand function.

Finally, in the agricultural economics literature, Zimmermann and Heckelei (2012) study the state dependence effect of the choice of farm size. They combine micro data on transitions with aggregate data at country level in order to get the best prediction at the country level. First, the transition probabilities are estimated using a minimal distance estimator where they must minimize distance with respect to both the transition patterns observed in the household level data and the market shares of the market level data. Second, they assume a multinomial logit function and estimate demand parameters using the transition probabilities estimated in the first place. Their model does not allow for unobserved individual heterogeneity, which, as I have mentioned earlier, is important in order to potentially identify structural state dependence.

The chapter is constructed as follows. In section 2.2, the data sets used are described. In section 2.3 the model is detailed. The estimation strategy is described in section 2.4. The results of the empirical application are discussed in section 2.5. Finally, I conclude in section 2.6.

2.2 Data

The market level data comes from the Nielsen Retail Scan dataset. These consist of weekly sales data at the Unique Product Code (hereafter UPC) recorded by retail stores. The household level data comes from the Nielsen HomeScan Panel. The dataset is based on information retrieved from the households who scan their purchases after each shopping trip. Both datasets are available for years 2005-2011. For this application, the latest year, year 2011 is used. Both datasets cover 205 designated market areas, which correspond to cities and their surroundings as defined by Nielsen.

For this study, two Midwest designated market areas are selected. Let us call these "City A" and "City B". My goal is to estimate a demand function for City A that accounts for consumer loyalty. However, the household sample of City A consists of 28 members, therefore I must rely on the market level data. For City B both the market level dataset and the household panel (with 1,734 households) are available. This setup is therefore matching the needs of the model described in this paper. In the first step, the loyalty parameters are estimated for the household panel of City B. Then the second step can be applied to City A, to retrieve demand parameter estimates which were computed using market level data but do account for loyalty.

I choose to estimate the model on the yoghurt product category. Yoghurt is a food obtained by bacterial fermentation of milk. Yoghurt has become increasingly popular in the US on the last two decades, yearly consumption per capita moving from 4.9 pounds in 1994 to 14.9 pounds in 2013[8]. The products are widely available, sold in the refrigerated area of most types of stores (from local shops to hypermarkets). Hence, this product category should ensure a sufficient sample size for the household level dataset. Yoghurt is diversified along various dimensions. Most yoghurt brands' product range offers a large variety of flavours (fruits, chocolate, vanilla, among many),

[8] Source: USDA Economic Research Services www.ers.usda.gov/data-products/dairy-data.aspx

milk fat levels (fat free, light, whole milk) and types (regular or greek). This product assortment should therefore be suitable to observe different types of household purchase history patterns in the data. I now describe each dataset in detail.

2.2.1 The Market Level Dataset: City A

The level of observation is a UPC store week. The dataset consists of 511 yoghurt UPCs observed in 98 stores over all 53 weeks of year 2011. The following information is recorded: the price, the total quantities sold, the median price for the UPC on that week, the quantity of yoghurt (expressed in OZ), the yoghurt firm and brand information, and yoghurt label information (fat free, light, regular, natural, Greek, organic, original and flavour).

To keep the model tractable, I aggregate the UPCs based on the producing firm brand, for the 5 largest firms and generics: Dannon, Fage, Stonyfield, The Greek Gods, Yoplait and Private Label (i.e. generics)[9]. The remaining products are classified as "Others". The observations are aggregated at the firm-week level. The market shares are computed based on quantities sold, and average prices are expressed in terms of USD per OZ. As can be seen in Table 2.1 the market is dominated by Dannon, Private Label and Yoplait, the three of which account, on average, for 90% of the market. Table 2.2 contains average weekly prices. One can see that Dannon and Yoplait price, on average, similarly at 0.11 USD per OZ. Fage yoghurt is, on average, twice more expensive.

Table 2.1: City A Retail Data, Market Shares

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Weeks**
Dannon	0.241	0.033	0.187	0.327	53
Fage	0.018	0.006	0.009	0.035	53
Stonyfield	0.019	0.003	0.014	0.026	53
The Greek Gods	0.014	0.002	0.011	0.024	53
Yoplait	0.355	0.044	0.291	0.465	53
Others	0.042	0.024	0.009	0.101	53
Private Label	0.311	0.021	0.271	0.355	53

Discounts are a common practice in the yoghurt industry. Table 2.3 shows discount frequency

[9] Private Label are unfortunately not separately identified by retailers in the dataset, therefore they must be proxied as one firm in this study.

Table 2.2: City A Retail Data, Prices

Variable	Mean	Std. Dev.	Min.	Max.	Weeks
Dannon	0.115	0.009	0.100	0.133	53
Fage	0.236	0.021	0.184	0.268	53
Stonyfield	0.165	0.013	0.134	0.200	53
The Greek Gods	0.130	0.006	0.115	0.140	53
Yoplait	0.109	0.007	0.094	0.121	53
Others	0.194	0.010	0.163	0.208	53
Private Label	0.072	0.002	0.063	0.075	53

and magnitude where a "discount"is defined as a price at least 5% lower than the previous week's price[10]. The first column contains the share of weeks where such discounts are observed. Firms seem to have different strategies. Fage products are discounted on 17% of the weeks. Private Label products are not surprisingly subject to discounts in 2% of the weeks only. Dannon is twice more frequently on discount than Yoplait. Finally, the last column shows the average size of the discount, conditional on the fact that the price was at least 5% lower than on the previous week. The differences among firms are less striking, their average discounts varying from 6 to 10%.

Table 2.3: City A Retail Data, Discount Frequency and Size

Firm	Share of Weeks On Discount	Size of Discount in Pct.
Dannon	0.17	0.07
Fage	0.21	0.09
Stonyfield	0.24	0.10
The Greek Gods	0.08	0.06
Yoplait	0.08	0.07
Others	0.08	0.08
Private Label	0.02	0.07

The product characteristics variables which relate to yoghurt label information are defined in the following way. The level of observation (a brand) is obtained by aggregating over UPCs. Hence, I create a variable which measures the share of UPCs labelled as (for example) fat free in the firm's product portfolio in that week. This is applied to the product characteristics light, regular, natural, Greek, organic and original as well. Regarding the flavours, I create a variable that counts the total number of flavours available in a firm's product portfolio in a given week. Table 2.4 provides descriptive statistics for these product characteristics. On average 40% of the product portfolio of a firm is fat free, while a firm in a given week carries on average 22.7 different flavors.

[10] The 5% threshold is a commonly used value in the literature, see for example Dube, Hitsch and Rossi (2010)

Table 2.4: City A Product Characteristics

Variable	Mean	Std. Dev.	Min.	Max.	Num. Obs.
Share of Fat Free Products	0.411	0.220	0.083	0.857	371
Share of Light Products	0.358	0.334	0	0.932	371
Share of Regular Products	0.208	0.311	0	0.917	371
Share of Natural Labelled Products	0.047	0.068	0	0.227	371
Share of Greek Labelled Products	0.467	0.397	0	1	371
Share of Organic Products	0.189	0.342	0	1	371
Share of Original Labelled Products	0.045	0.093	0	0.317	371
Number of Flavors	22.728	18.173	2	66	371

2.2.2 The Household Level Dataset: City B

In City B, the complete purchase history for each household is observed. The level of observation is a household UPC week. The dataset consists of 1,734 households observed on 67,906 weeks with a shopping trip and purchasing from 1,219 different UPCs. I aggregate this dataset at the firm household week level. Some caveats are in order. First, I exclude households purchasing more than one brand in a given week, in order to avoid the need to model basket purchases. This reduces the sample by 817 households. I also exclude households with less than 3 yoghurt purchases in the year, to ensure that a state variable is observed for each household. This results in losing 295 additional households. The resulting dataset consists of 622 households who are observed for a total of 24,664 weeks with shopping trips, buying yoghurt from 8 different firms[11]. The market shares, prices, and firm product portfolio variables are available in the same format as described in the previous section. Regarding the households, demographic variables are available and include: income, number of members and number of children.

Table 2.5 provides an overview of the choices made and the prices paid by the households during each trip. It shows that 80% of the trips do not include a yoghurt purchase. Yoplait is the firm with most purchases, being chosen in 7.3% of the shopping trips. Dannon follows with 5.4% and Private Label with 3.9%. Regarding the prices, the relative positions of the firms compared to one another are similar to those of Table 2.2, Private Label being the cheapest, Dannon and Yoplait having very similar mean prices. But in absolute terms, the prices of this household sample are below those of the market level data of City A for all brands (except Greek Gods, but

[11] Note that compared to City A, one additional firm is recorded, Mountain High. The reason being that its share of trips is relevant on this market. This illustrates a strength of the proposed framework in that differences between the product (here firm) sets of steps 1 and 2 does not limit the application of the model.

the number of observations for this brand is limited in the household sample). This was expected, as we are comparing the prices of realized purchases to the weekly published prices. Thus this difference suggests that consumers are on average attentive to price variations. Finally, the set of product characteristics is identical to the ones described in Table 2.4, and the corresponding table of descriptive statistics for City B can be found in the appendix in Table 2.A.2.

Table 2.5: City B Shopping Trips

Firm	**Average Price (USD/OZ)**	**Share of Trips**
Dannon	0.098	0.054
Fage	0.200	0.006
Mountain High	0.069	0.003
Stonyfield	0.168	0.005
The Greek Gods	0.195	0.001
Yoplait	0.097	0.073
Others	0.141	0.026
Private Label	0.070	0.039
No Purchase	-	0.793
Number of households		622
Avg. trips per household		39
Avg. purchases per household		8

Similar to most household scanner datasets, I do not have access to the prices faced by the household for the brands the household did not purchase. However, for the estimation of the maximum likelihood model of step 1, the prices for these alternative choices are required. The missing prices are proxied as follows. First, the data is merged at the store level between the household and the market level datasets, as some stores have records in both datasets. For unmatched products, missing prices are matched at the store level at the average prices paid by other households in the dataset. If not successful, missing prices are filled via interpolation if the prices for the weeks before and after are both available at the store of interest. If not successful, but the price of the week before (or after) is available at the store of interest, then this price is used as a proxy. Finally, if after these four steps the price is still missing, the average retail price over all shops in the market level dataset is used.

The key demographic characteristics of the sample are given in Table 2.6. The top panel contains information about the income of the households. The twelve income bins available in the dataset are aggregated into three bins. As can be seen, more than half of the households have a yearly income above 50 thousands USD per year. The middle panel shows that the sample is almost evenly split in terms of the number of household members. Finally, the bottom panel reveals that 80% of the households do not have children under 18 years old.

Table 2.6: City B Household Demographics

Category	Share
Less than 35k USD income	15.76
Income between 35-50k USD	21.38
More than 50k USD income	62.86
One hh member	28.25
Two members	39.33
Three or more member	32.42
No children	79.29
One child	4.17
Two children	3.85
Three children	8.67
Four or more children	4.01

The key feature of interest in the household level dataset is that purchase histories are observed at the individual level. In order to model state dependence and potentially identify loyalty, the purchase of today can be compared to past purchases. As mentioned in section 2.1, I follow a strand of the literature (see for example Erdem (1996), Dube, Hitsch and Rossi (2010)) by assuming that loyalty is measured by the last period's purchase. In order to compute descriptive statistics that are informative about loyalty I classify observed transitions (from one purchase to the other) as follows. If the household has purchased the same product as in the previous period, then the purchase is labelled as loyal, otherwise the purchase is not labelled as loyal.

The left panel of Figure 2.1 shows how the share of loyal purchases varies across households. While 33% of the households are always purchasing the same brand, there is 40% of the households whose share of loyal purchases is equal to or less than half. The right panel of Figure 2.1 shows the distribution of households with respect to the number of different firms from which they have been buying yoghurt throughout the year. While one third of the households purchase yoghurt from the same firm, 40% buy from two different firms and the remaining 30% buy from three or more different firms. These graphs suggest heterogeneity in households' purchase patterns.

To understand what variation in the data allows the identification of loyalty from that of individual heterogeneity, Figure 2.2 collects three examples of households showing very different purchase patterns. The dots show the yoghurt brands purchased over the year, and the labels show the ratio of purchase price to the yearly average price of the brand. On the first graph household

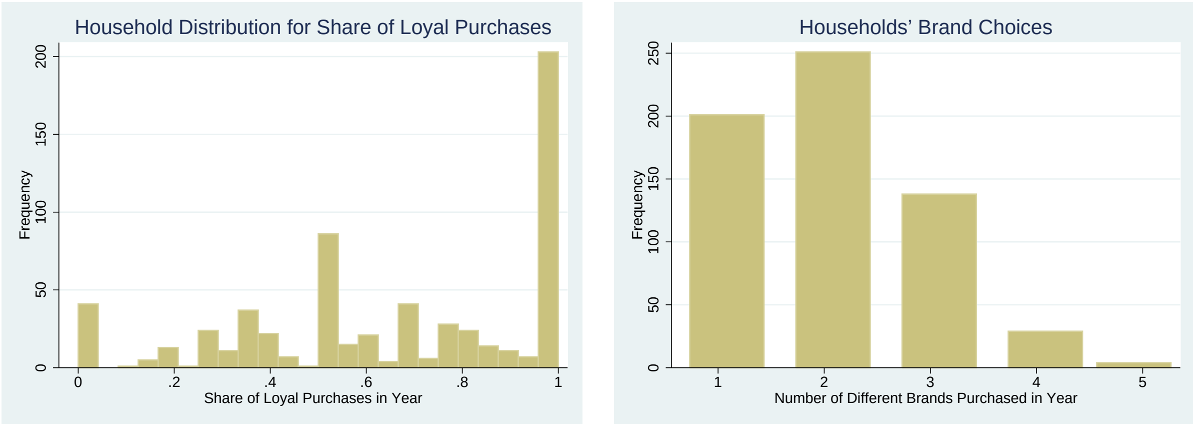


Figure 2.1: Heterogeneity in Households' Yearly Purchases

1 can be observed to always purchase the same brand. The price ratio is in almost half of the cases above 90%. Thus, this consumer's choice seems not to be influenced by prices, and rather be dominated by unobserved tastes for non-price product attributes. The second graph shows household 2 switching from one brand to the other, almost never consuming the same product in two consecutive periods. The price ratio is on average very low, moreover it is above 90% in only 2 of the 13 observed purchases. This seems to be a household who's choice is less affected by past purchases or strong brand preferences, and is likely to be driven by low prices. Finally, the third graph shows household 3 that switches between various products, but its behaviour does show some state dependence as a few repurchases are made before switching to another brand. Out of the 6 observed brand switches, 4 are initiated by a price ratio below 80%. This pattern seems to be potentially driven by a loyal behaviour: past purchase influence choice today, however, there is a switch to different products over time, potentially initiated by low prices. These are the different behaviours that the structural model described in the following section will allow to reproduce.

2.2.3 Approximation of the Outside Good

For each step of the model, an outside good has to be defined. Given the detailed household scanner data available, I define the outside good in the first step of the model as the decision of

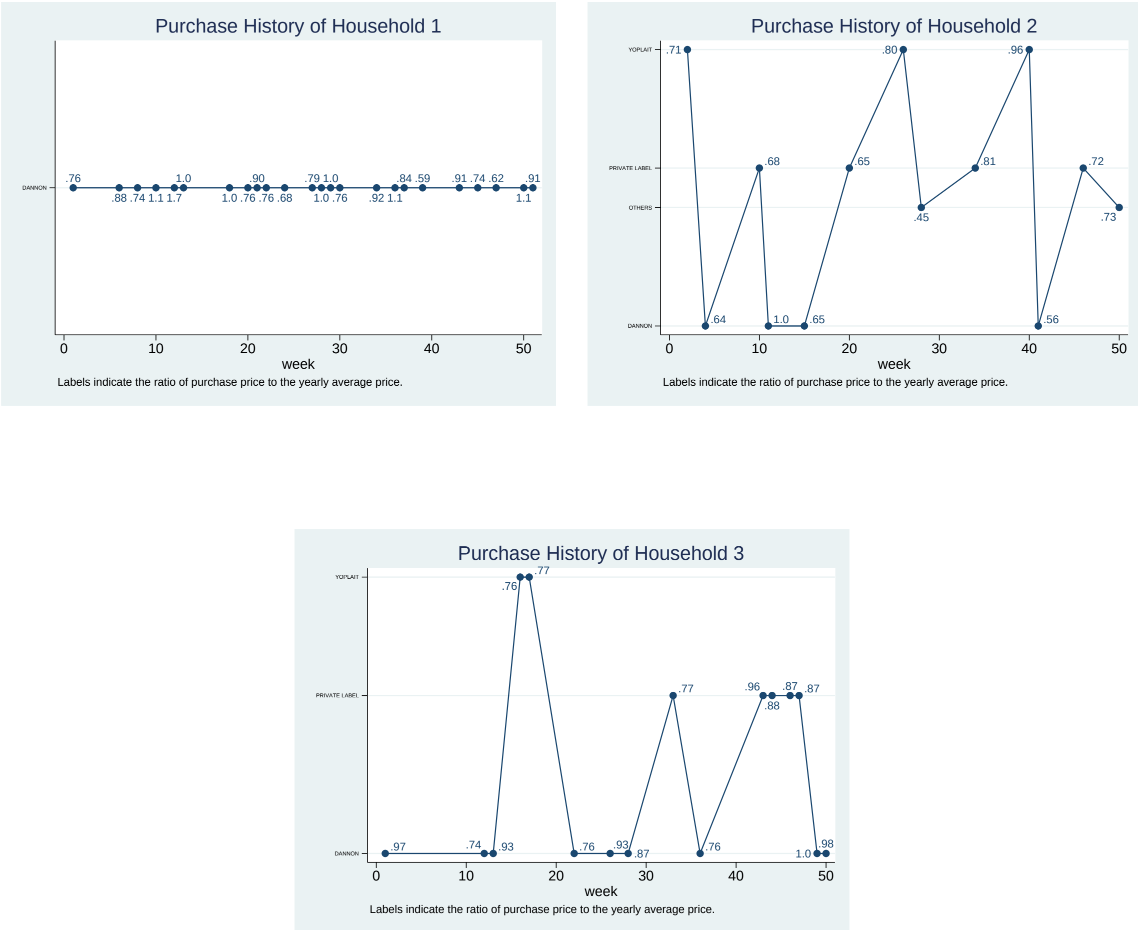


Figure 2.2: Three Examples of Purchase Histories

not purchasing any yoghurt during a shopping trip[12]. Over the whole household level sample, this leads to an average outside good share of 80%. Under the assumption that the sample is representative of the general yoghurt purchasing behaviour, this number is used to compute the outside good of the second step of the model. The market size there is thus proxied by taking the average total sales per period, and augmenting it by 80%. Then the outside good is obtained for each period as the market size minus the total sales in that period.

[12] One might be concerned that households stockpile yoghurt which is not accounted for by the model. However stockpiling affects the quantity of yoghurt purchased, while this paper focuses on the estimation of loyalty and its effect on subsequent brand choices. Therefore, ignoring stockpiling should not affect the results. For contributions related to stockpiling see Hendel and Nevo (2006) or Gordon and Sun (2015).

2.3 Model

The model is estimated in two steps. In a first step, a micro model is defined at the household level, where the dataset is only available for a subset of households. This allows the identification of state dependence as a heterogeneous coefficient, thereby distinguishing loyal and non-loyal households. In the second step, a market level demand function that accounts for state dependence is estimated on all markets. The market level model is based on the pair of moments following BLP: the predicted market shares must match observed ones and predicted unobserved product characteristics must be orthogonal to instruments. To augment the model with the state dependence term, three modifications are implemented. The state dependence coefficients are taken from the first step estimation and treated as data. Moreover, a proxy for the state variable (which otherwise is not defined in a market level model) is introduced, based on forward iteration of the predicted choice probabilities. Finally, a moment condition ensures that the share of households observed to be loyal in the micro data is matched by the predictions of the market level model.

The second step model is estimated in a GMM framework. The resulting demand function is valid for all markets (not just those where micro data was available), and accounts for loyal/non-loyal behaviour of households. Thus, the fit of the demand function for all markets should be improved.

The setup of the environment of the two models can be summarized as follows. For both micro and market level models, information is available on periods $t = 1, .., T$. The markets for which market level data is available are denoted by $m = 1, .., M$. The markets for which household level data is available as well are a subset $M^{micro} \subset M$.

2.3.1 First Step: Modelling the Household's Choice

The micro model is a static discrete choice problem. The household must decide, in each period, which product to purchase. In particular, the model accounts for the fact that households often buy the same product repeatedly, by allowing the last period's purchase to affect contemporaneous utility. As discussed in the introduction, households are assumed to be myopic.

Let us first describe the period utility of household h, with price coefficient of type r, loyalty type q, of consuming product j at time t given his previous period consumption choice z_{ht} as $U_{hrq}(j_t, z_{ht}; \theta)$, where θ is a vector containing the coefficients to be estimated. The utility provided by choosing product j can be split in two parts: the part that is identical across households δ_{jt} and the part that is household and type specific μ_{hrqjt}.

$$\begin{aligned} U_{hrq}(j_t, z_{ht}; \theta) &= \delta_{jt} + \mu_{hrqjt} \\ \text{where} \quad \delta_{jt} &= \gamma X_{jt} + \alpha P_{jt} + \xi_{jt} \\ \text{and} \quad \mu_{hrqjt} &= \alpha_r P_{jt} + \gamma_r X_{jt} + \kappa_q I_{j_t = z_{ht}} + \epsilon_{hrqjt}. \end{aligned} \tag{2.1}$$

The δ_{jt} term thus contains the base utility effect of: the product characteristics X_{jt} which have a base effect of γ, the base effect of the price of the product P_{jt}, α, and finally the unobserved product characteristics which are captured by ξ_{jt}.

The μ_{hrqjt} term contains the household and type specific part of the utility. It allows for heterogeneity along three dimensions. First, the coefficient κ_q allows the modeling of unobserved heterogeneity of household tastes for the last purchase z_{ht}. Second, the r-indexed coefficients allow the modeling of unobserved taste heterogeneity for the price α_r and for the other product characteristics γ_r. These coefficients measure deviations from the coefficients of the base utility γ, α. Finally, the indicator function on the state variable $I_{j_t = z_{ht}}$ accounts for the observed heterogeneity among households: the purchase that household h made in the previous period. The indicator function equals 1 if the product chosen in t, j_t, is identical to the state variable z_{ht} (i.e. the product chosen in $t-1$, j_{t-1}), otherwise the indicator function equals 0. The ϵ_{hrqjt} captures idiosyncratic differences among consumers' tastes. Following the literature (see for example BLP), I introduce the outside good, $j = 0$, to account for the decision of a household not to purchase any yoghurt. The utility of the outside good is normalized to depend solely on the idiosyncratic error term $u_{hrq0t} = \epsilon_{hrq0t}$.

The unobserved heterogeneity of households' tastes is modelled differently than the methodology described in BLP. These authors model random coefficients by assuming normally distributed tastes of the households. I follow Eizenberg and Salvo (2014) who use the modification suggested in Berry, Carnall and Spiller (1996): the tastes of the households are assumed to have a discrete distribution. Therefore, it is not necessary, as opposed to BLP, to draw random tastes and integrate over the draws to obtain the predicted market shares. Instead, a fixed number of household types, each with a set of coefficients and a weight is assumed. Thus the expected choice probabilities are obtained by computing the weighted sum over the types.

For the product characteristics coefficients, I define each type $r = 1, .., R$ to have a set of parameters γ_r, α_r[13] and a weight in the population w_r (the weights sum up to one). The types for the tastes over the state variable , κ_q, are modelled similarly. The index q models two types $Q = 2$. Loyal households, who get additional utility from purchasing the same brand as in last period are expected to have a positive κ_q. Non-loyal households who do not get additional utility from consuming the same brand as in last period, have their κ_q normalized to zero[14]. Alike the r types, each type q has a weight w_q in the population.

Given these R and Q types, the probability for household h to belong to type indexed rq, with coefficients $\alpha_r, \gamma_r, \kappa_q$ is thus $w_r w_q$. The parameters γ_r,α_r,κ_q and the weights w_r, w_q are estimated for all $r = 1, .., R$ and all $q = 1, .., Q$. Except w_R and w_Q, since these weights are given by the identities $w_R = 1 - \sum_{r=1}^{R-1} w_r$ and $w_Q = 1 - \sum_{q=1}^{Q-1} w_q$.

The distinction between the coefficient indices r and q is important. Indeed, κ_q and w_q can only be identified in the micro data available for the M^{micro} markets. They will therefore be estimated in the first step, and treated as data in the second step. The other random coefficients and weights $(\alpha_r, \gamma_r, w_r)$ however, are re-estimated in the second step, to benefit from the additional information stemming from the other markets available in the market level data M.

[13] For example with $R = 2$, the type-specific price coefficients $\alpha_1 = -1$ and $\alpha_2 = -3$ model two types of households: one with low and one with high sensitivity to prices.

[14] Modelling a third type of household, who would be variety seeking and would have a negative κ_q is of course a possibility. However, it is not in the focus of this paper and did not seem to be of relevance in the dataset used for the empirical application. Therefore, it is omitted in this model.

Assuming extreme value type I distributed error terms, the conditional probability for household h, of type rq with state variable z_{ht} to purchase product j in period t is:

$$s_{hrqjt}(z_{ht};\theta) = \frac{exp(\delta_{jt} + \alpha_r P_{jt} + \gamma_r X_{jt} + \kappa_q I_{j_t = z_{ht}})}{1 + \sum_j exp(\delta_{jt} + \alpha_r P_{jt} + \gamma_r X_{jt} + \kappa_q I_{j_t = z_{ht}})}.$$

2.3.2 Second Step: Modelling the Market Shares While Accounting for Sate Dependence

In the second step, a demand function at the market level is estimated on all markets $m = 1, .., M$. For the model to remain consistent with the first step, it is important for the utility function defined here to be as close as possible to the one used on the micro data, see equation (2.1). Therefore, the specification is almost identical, with two differences. First, the state dependence parameters κ_q and w_q are not estimated but given from the first step and treated as data. Second, the state variables are proxied using forward iteration over predicted choice probabilities.

Note that the following description of the model will be true for all M markets. However, for the sake of clarity, the m subscript will be omitted. The variables which vary across markets are: observed market shares S_{mjt}, predicted market shares $\hat{s}_{mjt}$ and prices P_{mjt}. Note that since the model estimates deep parameters and only the loyalty parameters are imported to the second step, the product sets across markets are allowed to differ.

In order to emphasize that the loyalty coefficients are treated as data, the coefficients estimated in the second step are denoted by θ^{-}. Hence, for a household of type rq the utility of choosing product j , with state variable $\tilde{z}_{rqt}$, based on the market level data (the market shares of all $j = 1, .., J$ products S_{jt}, the product characteristics X_{jt}, and the prices P_{jt}) can be written as:

$$U_{rq}(j_t, \tilde{z}_{rqt}; \theta^{-}) = \gamma X_{jt} + \alpha P_{jt} + \xi_{jt} + \alpha_r P_{jt} + \gamma_r X_{jt} + \hat{\kappa}_q I_{j_t = \tilde{z}_{rqt}} + \epsilon_{rqjt}.$$

Similar to the first step of the model, the household has an outside option: that of not purchasing yoghurt. The utility from the outside option is set as $U_{rq0t} = \epsilon_{rq0t}$.

Assuming again extreme value type I distributed error terms, the probability an individual of type rq with state variable $\tilde{z}_{rqt}$ purchases product j can be written as:

$$s_{rqjt}(\tilde{z}_{rqt};\theta^{-}) = \frac{exp(\delta_{jt} + \alpha_r P_{jt} + \gamma_r X_{jt} + \hat{\kappa}_q I_{jt=\tilde{z}_{rqt}})}{1 + \sum_{j'} exp(\delta_{j't} + \alpha_r P_{j't} + \gamma_r X_{j't} + \hat{\kappa}_q I_{j'_t=\tilde{z}_{rqt}})}.$$

As mentioned one difference with the first step model is that $\hat{\kappa}_q$ is data. The second difference concerns the state variable z_{rqt}. In the market level data, there is no information available on each type of agent's past purchase. Therefore, the previous period's purchase $\tilde{z}_{rqt}$ is proxied in the following way. Following Eizenberg and Salvo (2014), the probability for a certain product to be a state variable $\tilde{z}_{rqt} = j'$ for agent type rq is $\hat{s}_{rqj't-1}$, i.e. the choice probability predicted for the previous period. Thus, the choice probability unconditional on the state variable for agents of type rq for the current period t is obtained as follows:

$$s_{rqjt}(\theta^{-}) = \sum_{j'=1}^{J} \hat{s}_{rqj't-1}\hat{s}_{rqjt}(\tilde{z}_{rqt} = j';\theta).$$

Therefore, for a given guess of the initial conditions $\hat{s}_{rqj'0}$ the unconditional choice probabilities can be iterated forward for all periods. Finally, the predicted market share of product j in period t, is obtained by computing the weighted sum over the RQ different types of agents:

$$s_{jt}(\theta^{-}, W^{-}) = \sum_{r}\sum_{q} w_r \hat{w}_q \hat{s}_{rqjt}(\theta)$$

where W^{-} is the stacked vector of w_r for all r (since the loyalty weights w_q are not estimated).

2.4 Estimation Strategy

2.4.1 First Step: Estimating the State Dependence Parameters

The log-likelihood function is constructed for all households $h = 1, .., H$ observed on the M^{micro} markets using the conditional choice probabilities $s_{hrqjt}(z_{ht};\theta)$ described above, averaging over the RQ different household types and taking the joint probability over all periods T:

$$L(\theta, W) = \sum_{h=1}^{H} log\Big(\sum_{r=1}^{R}\sum_{q=1}^{Q} w_r w_q \prod_{t=1}^{T}\prod_{j=1}^{J} s_{hrqjt}(z_{ht};\theta)^{c^{\star}_{hjt}}\Big)$$

where $c^{\star}_{hjt}$ denotes the dummy variable of the observed choice of consumer h in period t for product j, and W is the stacked vector of w_r,w_q for all r, q.

Note that at this stage, the unobserved product heterogeneity ξ_{jt} has not yet been estimated. It is computed via a contraction mapping on the moment of predicted and observed market shares (as computed from the panel data)[15]. The parameters of the model are then estimated using a standard maximum likelihood procedure[16].

$$Max_{\theta,W} \quad L(\theta, W)$$

This gives us the estimates of the parameters of interest: the coefficients and weights corresponding to the loyal and non-loyal consumers[17]. The estimates are labelled as $\hat{\kappa}_q$,$\hat{w}_q$ and treated as data during the estimation of the second step.

2.4.2 Second Step: Estimating a Market Level Model Which Accounts for State Dependence

The estimation of the market level model follows closely the estimation strategy developed in BLP. The authors use two sets of moment conditions. The first moment accounts for the fact that the estimated market shares $\hat{s}_{jt}(\theta^{-})$ have to match the observed ones S_{jt}. The second moment condition is based on the orthogonality of the instruments (define them as IV_{jt}) with the unobserved product characteristics ξ_{jt}.

The first moment condition is used to compute the mean utility function δ_{jt} using the contrac-

[15] For more details see for example Train (2009).
[16] Derivatives of the first step maximum likelihood can be found in appendix 2.A.3.
[17] Note that the other estimated parameters are not used in the second step, so as to use as little information as possible from the markets of the first step estimation.

tion mapping technique described in BLP, with a slight difference. Here, the contraction mapping is computed separately for each period t (instead of being computed on all periods simultaneously in the original approach suggested by BLP). This breakdown is necessary since the state variables are computed via forward iteration. Thus, in period one, where the state variables are available through the initial conditions' assumption, the predicted market shares can be computed, and the $\hat{\delta}_{j1}$ obtained via contraction mapping over the moment condition: $\hat{s}_{j1}(\theta^-, W^-) = S_{j1}$.

Then one can move on to period 2 and proceed in the same way. Finally, once $\hat{\delta}_{jt}$ has been computed for all t, the vector of unobserved characteristics $\hat{\xi}_{jt}$ can be retrieved using the identity,

$$\hat{\xi}_{jt}(\theta^-, W^-) = \hat{\delta}_{jt} - \gamma X_{jt} - \alpha P_{jt}.$$

The procedure is applied to all markets $m = 1, .., M$ and the second moment is constructed as,

$$G_{IV}(\theta^-, W^-) = IV'\hat{\xi}(\theta^-, W^-)$$

where $\hat{\xi}(\theta^-, W^-)$ and IV are the stacked vector of the corresponding variables $\hat{\xi}_{mjt}(\theta^-, W^-)$ and IV_{mjt} (respectively). An additional moment condition is added to ensure that the market level model correctly mimics the behaviour of loyal/non-loyal households. The approach is in the spirit of Petrin (2002), who adds a "micromoment "condition to the above described BLP framework which helps the identification of the correlation between demographics and consumer tastes for product characteristics. In this paper, my goal is to account for the loyal/non-loyal behaviour observed in the micro data. Therefore, the micromoment I define states that the share of purchases predicted to be loyal in the market level model $\hat{P}^{marketlevel}$ must match the share of loyal purchases observed in the micro data $\hat{P}^{micro}$. The probabilities are defined as:

$$\hat{P}^{micro} = \frac{1}{T}\sum_t \sum_h \sum_j \frac{|c^\star_{hjt} = 1 \cap c^\star_{hjt-1} = 1|}{H}$$

$$\hat{P}^{marketlevel}(\theta^-, W^-) = \frac{1}{T}\sum_t \sum_r \sum_q w_r \hat{w}_q \sum_j \hat{s}_{rqjt-1} s_{rqjt}(\tilde{z}_t = j; (\theta^-, W^-))$$

The micromoment residual is thus defined as the difference between the two probabilities:

$$G_{Loyal}(\theta^-, W^-) = \hat{P}^{micro} - \hat{P}^{marketlevel}(\theta^-, W^-).$$

Finally, the objective function of the GMM for the second step of the model is built following Hansen (1982), and is minimized to recover the parameters θ^- and W^-.

$$Min_{\theta^-, W^-} \quad G'(\theta^-, W^-) \quad A \quad G(\theta^-, W^-) \tag{2.2}$$

where $G(\theta^-, W^-) = \begin{pmatrix} G_{IV}(\theta^-, W^-) \\ G_{Loyal}(\theta^-, W^-) \end{pmatrix}$ and $A = \begin{pmatrix} A_{IV} & 0 \\ 0 & A_{Loyal} \end{pmatrix}$ is a weight matrix[18].

As described in Hansen (1982) the consistent and efficient estimator is obtained by estimating the GMM model above (the second step) multiple times, iterating over the value of the weight matrix A. Based on the comment in Nevo (2000), I iterate the value of the weight matrix exactly once since there is no strong evidence that more iterations bring significant improvements in finite sample problems. First, the weight matrix A is set to identity, and parameter estimates $\hat{\theta}^-$ and $\hat{W}^-$ are retrieved. Using these preliminary consistent estimates, the proxy for the weight matrix is computed for A_{IV} (and similarly for A_{Loyal}) as follows :

$$\hat{A}_{IV} = \left(\frac{1}{MJT} \sum_{mjt} G_{IVmjt}(\hat{\theta}^-, \hat{W}^-) G'_{IVmjt}(\hat{\theta}^-, \hat{W}^-) - \bar{G}_{IV}(\hat{\theta}^-, \hat{W}^-) \bar{G}'_{IV}(\hat{\theta}^-, \hat{W}^-) \right)^{-1},$$

where $\quad \bar{G}_{IV}(\hat{\theta}^-, \hat{W}^-) = \frac{1}{MJT} \sum_{mjt} G_{IVjt}(\hat{\theta}^-, \hat{W}^-).$

The objective function described in equation (2.2) is then re-estimated using $\hat{A}$, to obtain the final estimates for $\hat{\theta}^-$ and $\hat{W}^-$.

The set of parameters resulting from the iteration of the second step GMM, $\hat{\theta}^-$ and $\hat{W}^-$, with the parameters $\hat{\kappa_Q}$ and $\hat{W}_Q$ taken along from the first step estimation build the demand function of the model: a market level demand function which accounts for consumer loyalty.

[18] The weight matrix is block-diagonal as the two sets of moments are based on two independent sampling processes.

2.4.3 Identification

In what follows, I discuss informally the identification of the parameters and the selection of instruments. I first discuss the identification of the parameters that appear in both the first and the second steps of the model: price and product characteristics. The market shares' correlation with the observed product characteristics allows the identification of the coefficients of the base utility (α and γ). In the empirical application, the product characteristics that measure the share of different sorts of yoghurt in the firm's portfolio vary across firms, but also over time. The prices differ across firms, and show significant volatility over time as well. Regarding the identification of the type specific coefficients (α_r and γ_r) and weights (w_r), the explanation follows the literature (see for example Berry, Carnall and Spiller (1996), BLP). Assume that in period t, there are 4 products which are identical, except regarding their prices. Two products are priced p_1 and two are priced p_2. Assume that they have equal market shares. Assume that in period t', one of the p_2 priced products has its price changed to p_1. In that case, a homogeneous logit estimation would predict for t' identical market shares for all four products. However, in reality (with heterogeneous price sensitivity of households) one would expect in t' the three products at price p_1 to be in stronger competition and therefore see their market shares decrease. These substitution patterns allow the identification of the type-specific coefficients.

I now turn to the identification of the loyalty parameters, which depends solely on variation in the micro data. In a given period, many different households' purchases are observed along the brand purchased in the previous period. Therefore, for given product prices and characteristics, the variation between the product chosen in the period and the state variable allows the identification of the loyalty parameters(κ_q and w_q).

Finally, I discuss the selection of instruments. The estimation procedure of the second step of the model is based on the mean independence condition of the unobserved product characteristics ξ and the instruments. The ξ's are unobserved to the econometrician but consumers do observe them. Thus, firms are likely to account for them in the prices they are setting. Therefore, the instruments need to be correlated with prices, and uncorrelated with ξ. As described in BLP, a possible solution consists of using variables which shift markups. As they note if many products are substitutable, the mark-up should be low, and vice versa. I follow these authors' suggestions: I use various combinations of the product characteristics, including the square, the cubic power and the sum of the product characteristics. Another type of instruments, suggested in Hausman (1996), is also used: the prices of the J goods on other markets. These should be correlated with the price on the market of interest as the products should stem from identical factories and therefore share same costs. Moreover, the concern (as discussed in Bresnahan (1997)) that such instruments might be correlated with unobserved product characteristics should be mild: in the empirical application presented in this paper, the product assortment is varying across markets hence the local competition should limit this correlation.

2.5 Results

The results of the paper are described in four subsections. First, estimation results are presented for the model using Monte Carlo generated data. The reader shall here be convinced that the model is performing well in a synthetic environment. Second, a few preliminary regressions are proposed to select explanatory variables and instruments. In the third subsection, the results of the model estimated on the yoghurt dataset described in section 2.2 are presented. Finally, I discuss and compare these findings.

2.5.1 Monte Carlo Results

The data for the Monte Carlo simulations is generated following the model defined in the previous section. Let us call City X, the city without household level data, and City Y the city where household level data is available. For each city, a specific set of 25 products is generated (and an outside good), with 3 product characteristics (constant over periods) and prices for each of the 50 periods.

I discuss the results of two different simulation setups. In the first one, the datasets of City X and City Y are generated with identical price coefficients. In the second setup, the true values for the price coefficients of the two cities differ.

I now describe briefly the data generation process in the case where price coefficients differ between City X and City Y. For City X, I generate 5000 households who make a choice in every period. I define two types of price sensitivities for households: one half of the households have a price coefficient of -.5, for the other half the price coefficient is set at -2. I model loyalty in a conservative way: one half of the households are set to be loyal with a coefficient of 3. The other half of the households is set as indifferent to past purchase with a normalized coefficient at 0. Regarding the three other product characteristics, households have identical coefficients with 1.5, 1.5 and 1.6. Hence, the population is made of 4 equally represented types. For each simulated household, random shocks (ϵ) are drawn and thus the product providing the highest utility is recorded as the choice in each period. Then these individual choices are aggregated and I obtain the market shares dataset. For City Y, purchase histories for 1000 households are simulated following the same procedure, with two differences. In City Y, the true price coefficients of households are set at -1 and -3 respectively. Moreover, the choices are not aggregated, the household purchase histories build the household level dataset for City Y[19].

The results of the estimation are benchmarked to a standard for demand estimation on market level data: the BLP model[20]. This model is similar to the model described in the second step. The coefficients on product characteristics and weights are defined in the same way. The only (but key) difference being that the loyalty term is here not modelled. Thus, the comparison to this benchmark will reveal the importance of accounting for loyalty.

Results are shown in Table 2.7 and Table 2.8. The means and confidence intervals are computed over the estimation results on 50 simulated Monte Carlo datasets. In the left panel, the

[19] Codes for the Monte Carlo processes are available from the author upon request.

[20] It should be mentioned that the benchmark is based on less information than the two step model since it is not using micro data. An alternative approach would be to use as a benchmark the model in Berry, Levinsohn and Pakes (1998) or in Petrin (2002) which exploit the limited available micro data using micromoments relating demographics to consumption choices. In that case, the present model could be modified to additionally account for identical micromoments.

first column shows results for the Maximum Likelihood Estimation for City Y. The second column shows results for the Modified BLP (with Loyalty) Estimation for City X. The third column shows results for the BLP (without loyalty) Estimation for City X. The right panel of the tables contains the true values of the parameters for each city.

Table 2.7 contains results for the case where the cities have identical price coefficients. The coefficients are best estimated via the Maximum Likelihood estimation. The focus of this paper is on the comparison of the Modified BLP and the Benchmark BLP. As it can be seen, modifying the BLP model to account for loyalty does bring an improvement. For all the Modified BLP coefficients and weights, the means are closer to the true values, and the confidence intervals are narrower. Finally, the goal of this paper is to improve the demand function fit. To measure the performance of the models, I compute the Mean Squared Error over market shares[21]. It reveals that the model described in this paper, the Modified BLP is performing on average 7.5% better than the Benchmark BLP (without loyalty). The improvement is quite significant.

Table 2.7: Simulation Results: Cities with Identical Price Coefficients

	Two Step Model		**Benchmark**	**True Values**
	Max. Like. City Y	Modified BLP City X	BLP City X	Both Cities
Price Type 1	-1.206	-1.126	-0.770	-1.000
	[-1.28,-1.13]	[-1.23,-1.03]	[-1.04,-0.50]	
Price Type 2	-3.188	-3.634	-3.984	-3.000
	[-3.26,-3.11]	[-3.89,-3.37]	[-4.88,-3.09]	
Loyalty	2.948	2.948†	-	3.000
	[2.89,3.01]	[2.89,3.01]	-	
Weight Type 1	0.499	0.507	0.435	0.500
	[0.49,0.50]	[0.47,0.54]	[0.39,0.47]	
Weight Loyal	0.480	0.480†	-	0.500
	[0.47,0.48]	[0.47,0.48]	-	
Prod. Char. 1	1.565	1.694	1.668	1.500
	[1.26,1.87]	[1.39,1.99]	[1.37,1.97]	
Prod. Char. 2	1.544	1.671	1.663	1.500
	[1.24,1.85]	[1.38,1.96]	[1.37,1.96]	
Prod. Char. 3	1.932	1.931	1.972	1.600
	[1.73,2.13]	[1.74,2.12]	[1.79,2.15]	
Constant	0.224	0.271	0.437	0.100
	[0.13,0.31]	[0.13,0.41]	[0.28,0.59]	
MSE	1.483	1.502	1.629	-
	[1.482,1.483]	[1.502,1.503]	[1.628,1.629]	

Means computed over the estimation results of 50 randomly generated datasets. The 95% confidence intervals are shown in squared brackets. † indicate that loyalty parameters are imported from the first step Maximum Likelihood estimation.

Table 2.8 contains results for the case where the cities have different price coefficients. Compared to the previous case, the three estimations rank in the same way. However, the Modified BLP is now only slightly closer to the true value than the classical BLP. Except for the price

21 These predicted market shares are computed by setting the unobserved product characteristics ξ to zero (a common procedure for BLP models)

coefficient of type 2 households, where the mean is one unit closer to the true value. The MSE over market shares confirms that the advantage of the modified BLP has diminished: it provides only a 2% improvement in MSE compared to the classical BLP. These results therefore seem to indicate that the model is sensitive to differences in true parameters between the markets on which it is estimated. The more different the markets are, the smaller will be the advantage of the two-step model presented in this paper.

Table 2.8: Simulation Results: Cities with Different Price Coefficients

	Two Step Model		**Benchmark**	**True Values**	
	Max. Like. City Y	Modified BLP City X	BLP City X	City Y	City X
Price Type 1	-1.179	-0.193	0.197	-1.000	-0.500
	[-1.31,-1.05]	[-0.41,0.03]	[-0.45,0.84]		
Price Type 2	-3.158	-2.581	-3.520	-3.000	-2.000
	[-3.29,-3.03]	[-2.84,-2.32]	[-5.12,-1.92]		
Loyalty	2.943	2.943†	-	3.000	3.000
	[2.94,2.95]	[2.94,2.95]	-		
Weight Type 1	0.499	0.455	0.479	0.500	0.500
	[0.49,0.50]	[0.41,0.50]	[0.39,0.56]		
Weight Loyal	0.483	0.483†	-	0.500	0.500
	[0.48,0.49]	[0.48,0.49]	-		
Prod. Char. 1	1.586	1.415	1.411	1.500	1.500
	[0.91,2.26]	[0.97,1.86]	[0.99,1.83]		
Prod. Char. 2	1.564	1.413	1.395	1.500	1.500
	[0.88,2.25]	[0.96,1.86]	[0.97,1.82]		
Prod. Char. 3	1.790	1.919	1.924	1.600	1.600
	[1.42,2.16]	[1.60,2.23]	[1.61,2.22]		
Constant	0.252	0.712	0.309	0.100	0.100
	[0.08,0.43]	[0.43,0.99]	[0.09,0.52]		
MSE	1.599	1.891	1.926	-	-
	[1.598,1.600]	[1.890,1.891]	[1.926,1.927]		

Means computed over the estimation results of 50 randomly generated datasets. The 95% confidence intervals are shown in squared brackets. † indicate that loyalty parameters are imported from the first step Maximum Likelihood estimation.

2.5.2 Preliminary Analysis

Before presenting the estimation results of the structural model on cities A and B of the yoghurt purchases data described previously, I provide a short reduced form analysis in order to select dependent variables and instruments. I present four homogeneous logit specifications. This is an ordinary least squares regression model where the dependent variable is given by the difference between the logarithms of the market share and the outside good.

Tables 2.9 and 2.10 show the estimation results for City A and City B respectively. In the first column of each, all explanatory variables are listed. The following two columns contain estimation results for specifications based on the significant variables in column one. The price coefficient is negative and significant on both markets. It is almost twice larger in City B than in City A. The significance of the product characteristics varies depending on the cities as well as on the specifications. The variables that remain stable qualitatively and in terms of significance are the share of Natural and Greek products. Given the significance levels the set of variables in columns two and three are the ones on which the structural model is estimated.

Table 2.9: Homogenous Logit Estimates City A

Variables	OLS	OLS	OLS	IV
Price USD/OZ	-6.987***	-6.993***	-7.006***	-8.400***
	(1.091)	(1.104)	(1.099)	(1.411)
Fat Free	-0.299	-0.424		
	(0.496)	(0.497)		
Light	-0.199	-0.196		
	(0.578)	(0.583)		
Regular	-1.718**	-1.833***	-1.411***	-1.350***
	(0.666)	(0.671)	(0.480)	(0.474)
Natural	-4.281***	-4.355***	-4.398***	-4.346***
	(0.526)	(0.533)	(0.507)	(0.500)
Greek	1.088***	1.013***	0.966***	1.035***
	(0.329)	(0.332)	(0.302)	(0.301)
Organic	0.469			
	(1.339)			
Original	6.182***			
	(1.766)			
Flavors	5.561***	4.985***	5.040***	5.009***
	(0.755)	(0.741)	(0.728)	(0.717)
Constant	-4.133***	-3.837***	-4.179***	-3.956***
	(0.601)	(0.598)	(0.328)	(0.354)
Observations	371	371	371	371
R-squared	0.976	0.975	0.975	0.975
First Stage				
Adjusted R^2				.983
F-statistic				1307
Prob>F				0

Dependent variable: ratio of the log of market share and the log of the outside good market share. All regressions include firm dummies and a time trend. Standard errors in parentheses. Stars indicate coefficient significance level at 10% (*), 5% (**) and 1% (***)

The fourth column proposes an instrumental variables (IV) regression to account for the endogeneity of the price variable. The instruments typically used in the literature (following BLP), are combinations of the product characteristics and include the square, the cubic power and the sum of the product characteristics. In this application, these had very little correlation with the endogenous variable. Therefore another set of instruments was used following Hausman (1996), the prices of the goods on other retail markets. These perform much better and are the

Table 2.10: Homogenous Logit Estimates City B

Variables	OLS	OLS	OLS	IV
Price USD/OZ	-15.12***	-14.52***	-14.70***	-14.07***
	(0.912)	(0.919)	(0.925)	(1.577)
Fat Free	-0.873	-2.875***		
	(0.931)	(0.825)		
Light	-1.059**	-1.796***		
	(0.513)	(0.494)		
Regular	-1.341	-3.210***	-0.324	-0.389
	(0.962)	(0.876)	(0.337)	(0.357)
Natural	-6.442***	-7.323***	-6.865***	-6.846***
	(0.943)	(0.938)	(0.925)	(0.911)
Greek	1.572**	2.129***	1.879***	1.761***
	(0.705)	(0.700)	(0.496)	(0.544)
Organic	4.107***			
	(0.944)			
Original	-0.771			
	(1.200)			
Flavors	0.826**	0.903***	0.478	0.457
	(0.320)	(0.323)	(0.298)	(0.296)
Constant	-0.341	1.230**	-0.644*	-0.711*
	(0.700)	(0.607)	(0.362)	(0.382)
Observations	424	424	424	424
R-squared	0.989	0.988	0.988	0.988
First Stage				
Adjusted R^2				.983
F-statistic				1447
Prob>F				0

Dependent variable: ratio of the log of market share and the log of the outside good market share. All regressions include firm dummies and a time trend. Standard errors in parentheses. Stars indicate coefficient significance level at 10% (*), 5% (**) and 1% (***)

ones used for the IV regression presented in the fourth column. The results reveal that the instruments impact strongly the price coefficient in City A (Table 2.9), where it gets 14% larger in absolute terms. However, the price coefficient in City B (Table 2.10)is almost unaffected. The other variables keep the same magnitude and significance levels compared to the OLS estimates. These instruments are the ones used for the estimation of the structural model.

2.5.3 Results of the Structural Estimation

The structural model proposed in this paper is estimated using the explanatory variables and instruments selected in the previous subsection. Table 2.11 presents the results. The left panel contains the specification with six product characteristics, the right panel contains the one with four product characteristics. For each of them, the first column contains the estimation results of step 1, the Maximum Likelihood estimation over the household panel of City B. The second column contains the estimation results of step 2, the modified BLP model which accounts for loyalty estimated on market level data of City A. Finally, for comparison, the third column contains the benchmark model, the estimates of a BLP model with discrete types for City A [22].

For both specifications, each model has been estimated using 50 different random starting values, spanning a large range of potential values for the parameters. The Maximum Likelihood estimation of the first step has proved to be extremely robust, converging in all fifty cases to the same optimum. The second step GMM estimation has proved less stable: several local optima were found, mostly with one unique type of price coefficient and two price coefficient types in 5% of the runs. Table 2.11 presents the results of the runs which performed best in terms of the Mean Squared Error (hereafter MSE) measure over the market shares.

Let us first compare the results of the two specifications in the left and the right panel. The coefficient estimates are stable for each city. The Maximum Likelihood estimates converge exactly to the same point in both cases in terms of the price coefficients and of the weights. This is not the case for the coefficients of the other product characteristics. However, these coefficients are not of key interest since the goal of the first step Maximum Likelihood is to retrieve the loyalty parameters. The modified BLP estimates are stable across the specifications. The benchmark BLP estimates are also stable across specifications.

I now discuss the results for each of the models separately. The Maximum Likelihood model over household panel data identifies two groups of households with different price coefficients. The low price sensitive households have a price coefficient of -6.7 and are estimated to be 32% of the sample. The remaining part of the sample contains more price sensitive households with a price coefficient of -20.2. The share of loyal household types is estimated at 60% with a loyalty coefficient of 2.6, while the remaining share of households is non-loyal (with a normalized coefficient of 0).

The Benchmark BLP model over market data has converged to a unique type for the household price coefficient in City A. This type has a price coefficient of -8.4. This coefficient and those of the other product characteristics are in line with the homogenous logit results of Table 2.9.

[22] Because the BLP model with discrete distribution of tastes have converged to a unique type solution, the jacobian is multicolinear, therefore only the standard deviation of the active (non-zero estimated) coefficient is reported.

Finally, let us analyse the results of the second step of the model presented in this paper. The Modified BLP model has converged to a unique type for the household price coefficient in City A. This type has a price coefficient of -8 and is significant. The coefficients for the share of Fatfree and Light products are insignificant. The coefficients for the share of Regular Natural Greek and Flavour products are significant and their values are in line with the homogenous logit results of Table 2.9. This suggests that the addition of the loyalty parameter from the first step Maximum Likelihood estimation (while comparing to the results of the Benchmark BLP model) affects the estimation results very slightly, the price coefficient getting lower by 5% in absolute terms.

I now turn to the comparison of the performance of the Modified BLP to the Benchmark BLP in order to verify whether the importing of the loyalty parameters from the first step Maximum Likelihood has brought improvement in terms of fitting the market share data. The MSE on market shares is presented in the bottom panels of Table 2.11. In both specifications, the models rank similarly. The Maximum Likelihood on household level data has the lowest MSE. The Benchmark BLP on market level data performs 2% worse in terms of MSE. Finally, the Modified BLP with loyalty, the model described in this paper, performs 8% worse than the Benchmark BLP estimation.

2.5.4 Discussion of the Results

The performance of the model differs significantly whether estimated on the simulated datasets or the yoghurt data used in the empirical application. Estimated on simulated datasets, the Modified BLP performs 7.5% better than the Benchmark BLP if the price coefficients of the two markets are identical, and 2% better if the price coefficients differ. However, when estimated on the yoghurt data described in section 2.2, the Modified BLP performs 8% worse than the Benchmark BLP. I now provide some explanations for why this might occur and suggest potential solutions.

First, this paper's goal is to propose a demand estimation that accounts for loyalty behaviour for markets where household level data is not available. This is achieved by estimating the loyalty

Table 2.11: Estimation Results

	Specification 1			**Specification 2**		
	Two Step Model		**Benchmark**	**Two Step Model**		**Benchmark**
	Max. Like.	Modified BLP	BLP	Max. Like.	Modified BLP	BLP
	City B	City A	City A	City B	City A	City A
Price Type 1	-6.659***	-8.004***	-8.407***	-6.659***	-7.978***	-8.393***
	(0.012)	(1.534)	(1.532)	(0.012)	(1.508)	(1.507)
Price Type 2	-20.206***	-19.465	-19.466	-20.206***	-19.449	-19.468
	(0.012)	(-)	(-)	(0.012)	(-)	(-)
Loyalty	2.585***	2.585†	-	2.585***	2.585†	-
	(0.001)	(-)	(-)	(0.001)	(-)	(-)
Weight Type 1	0.322***	0.998	0.996	0.322***	1.000	0.996
	(0.001)	(-)	(-)	(0.001)	(-)	(-)
Weight Loyal	0.598***	0.598†	-	0.598***	0.598†	-
	(0.002)	(-)	(-)	(0.002)	(-)	(-)
Fatfree	2.140	-0.179	-0.400	-	-	-
	(2.188)	(0.440)	(0.461)	(-)	(-)	(-)
Light	-0.438	-0.010	-0.142	-	-	-
	(2.497)	(0.639)	(0.658)	(-)	(-)	(-)
Regular	0.911	-1.505**	-1.762**	-1.065	-1.306**	-1.351**
	(1.548)	(0.656)	(0.675)	(1.570)	(0.590)	(0.591)
Natural	-8.185***	-3.956***	-4.289***	-7.317***	-4.003***	-4.346***
	(1.867)	(0.749)	(0.776)	(1.894)	(0.691)	(0.719)
Greek	-4.009	1.014**	1.093**	-1.596	0.970**	1.035**
	(2.598)	(0.467)	(0.475)	(2.635)	(0.380)	(0.389)
Flavor	-1.065	4.610***	4.944***	-0.635	4.656***	5.010***
	(3.125)	(1.112)	(1.140)	(3.160)	(1.102)	(1.137)
MSE	0.058	0.064	0.059	0.057	0.064	0.060
Num. Obs.	24664	371	371	24664	371	371

All regressions include firm dummies, a time trend and a constant. Standard errors in parenthesis. Stars indicate coefficient significance level at 10% (*), 5% (**) and 1% (***). † indicate that loyalty parameters are imported from the first step Maximum Likelihood estimation.

parameters on another market's household data. Then these parameter estimates are imported to the market level model. The classical BLP model is modified to account for loyalty and is estimated on the market of interest. As discussed in the introduction, this strategy relies on the assumption that the two markets are quite similar in terms of the loyalty parameters: the coefficient in the utility function and the share of loyal-type households. Therefore, in case this assumption is strongly not met (for example one market has a large majority of loyal households while the other market is mostly populated by non-loyal households) importing the estimated loyalty parameter will likely lead to the model performing worse than the benchmark model (not accounting for loyalty). I must remind the reader here that this assumption is not verifiable prior to estimating the structural model, since the whole paper is designed to estimate a demand function which accounts for loyalty for a market where household data is not available. Thus, the fact that the model proposed in this paper performs worse than the Benchmark BLP estimation might just reveal that the markets used in the empirical application are populated with households having too different behaviours in terms of loyalty. This is probably the case of the empirical application of the current paper. For such cases, where a first attempt using household level data from a specific market does not prove to be successful, I suggest the following solution. This solution relies on the assumption that a larger set of markets with household level data is available. The researcher could then run the first step Maximum Likelihood estimation on each of these markets. This would provide a set of estimated values for the loyalty coefficients and weights. Then, the second step, the Modified BLP model, could be run with each element of this set and the MSE over market shares can be computed. Finally, those coefficients and weights leading to the best performing MSE over market shares could be retained as the preferred solution.

Another potential explanation for the performance of the model in the empirical application can be related to the assumption that the loyalty coefficient is identical across brands. Indeed, it could be that loyalty is brand-specific. In that case, for a given price change, the homogeneous loyalty coefficient in the model would be underestimating the share of consumers staying loyal to the product for the brand which has a strong loyalty coefficient and overestimating it for brands which do not. The solution to this problem can be implemented in the model: the κ_q parameters have to be defined as j specific. The results could then be compared to the current model and the Benchmark BLP. This modification is left for future research.

2.6 Conclusion

In this paper, I propose a model to account for consumer loyalty in market level demand functions when household level data is not available. This is achieved by exploiting household level data available for a subset of markets. The model is based on two steps. In a first step, a micro model is defined at the household level. Because the households' purchase histories are observed, the loyalty parameters can be estimated. In the second step, a market level demand function that accounts for state dependence is estimated on all markets. This model is similar to the BLP procedure, with three additions. The loyalty coefficients are taken from the first step estimation and treated as data. A proxy for the state variable (which is not defined in a market level model) is introduced, based on forward iteration of predicted choice probabilities. Finally, an extra moment condition is added, which ensures that the share of loyal purchases predicted by the market level model matches the one observed in the micro data.

The model's performance is compared to the BLP model (a widely used model for demand function estimation on market share data) and tested in two different environments. First, the model is estimated on Monte Carlo simulated datasets. The model performs 2% to 7.5% (in terms of MSE over market shares) better than the BLP model depending on how different the data generating processes are. These results prove that the model does improve the fit of market shares. Second, the model is estimated on yoghurt sales for two Midwestern cities based on Nielsen data. The results over the synthetic datasets are here unfortunately not confirmed. The model performs 8% worse than the benchmark BLP. I provide potential explanations for the failure of the model on this empirical application. This model requires the assumption that the driving forces of households' loyal/non-loyal behaviour have to be similar on the market with household data and the market to which the estimates are imported. However, this is an assumption that cannot be tested, since the model is designed to provide a solution to the lack of household data for the market of interest. Therefore, if the model fails to perform better than the benchmark BLP, it might be indicating that the market from which the loyalty parameters are imported is populated with households having markedly different characteristics from those of the market of interest. One solution would be in running the first step of the model on all markets where household level data is available. This would result in a set of estimated loyalty coefficients and weights. Then, the second step of the model could be run for each element of this set, and the element that would lead to the best performing estimation retained, resulting in a market level demand function that accounts for loyal/non-loyal behaviour. This approach can be of interest for future research.

References

1. Berry S., Levinsohn J. and Pakes A. (1995) "Automobile prices in market equilibrium", *Econometrica*, 63(4), 841-890.

2. Berry, S., Levinsohn, J. and Pakes, A. (1998) "Differentiated products demand systems from a combination of micro and macro data: The new car market", *National Bureau of*

Economic Research, w6481.

3. Berry S., Carnall M. and Spiller P. (1996) "Airline hubs: costs, markups and the implications of customer heterogeneity", *National Bureau of Economic Research*, Technical Report.

4. Bresnahan, T. F. (1997) "The apple-cinnamon cheerios war: Valuing new goods, identifying market power, and economic measurement", *The Economics of New Goods*, The University of Chicago Press, Chicago/London.

5. Chen Y. and Yang S. (2007) "Estimating disaggregate model using aggregate data via augmentation of individual choice", *Journal of Marketing Research*, 4 596-613.

6. Dube J.P., Hitsch G. and Rossi P. (2010) "State dependence and alternative explanations for consumer inertia", *The RAND Journal of Economics*, 41(3), 417-445.

7. Eizenberg A. and Salvo A. (2014) "The rise of fringe competitors in the wake of an emerging middle class: An empirical analysis", *American Economic Journal: Applied Economics*, 7(3), 85-122.

8. Erdem T. (1996) "A dynamic analysis of market structure based on panel data", *Marketing Science*, 15(4), 359-378.

9. Gordon B. and Sun B. (2015) "A dynamic model of rational addiction: Evaluating cigarette taxes", *Marketing Science*.

10. Guadagni P. and Little J. (1983) "A logit model of brand choice calibrated on scanner data", *Marketing science*, 2(3), 203-238.

11. Hansen L.P. (1982) "Large sample properties of generalized method of moments estimators", *Econometrica*, 50(4), 1029-1054.

12. Hausman J. (1996) "Valuation of new goods under perfect and imperfect competition", In *The Economics of New Goods*, University of Chicago Press, 207-248.

13. Heckman J. (1981) "Heterogeneity and state dependence", *Studies in Labor Markets*, University of Chicago Press, 91-140.

14. Hendel I. and Nevo A. (2006) "Measuring the implications of sales and consumer inventory behavior", *Econometrica*, 74(6), 1637-1673.

15. Horsky D., Pavlidis P. and Song M. (2012) "Incorporating state dependence in aggregate brand-level demand models", Technical report, Mimeo, University of Rochester.

16. Musalem A., Bradlow E. and Raju J. (2009) "Bayesian estimation of random coefficients choice models using aggregate data", *Journal of Applied Econometrics*, 24(3), 490-516.

17. Nevo A. (2000) "A practitioner's guide to estimation of random-coefficients logit models of demand", *Journal of Economics & Management Strategy*, 9(4), 513-548.

18. Newey W. K. and McFadden D. (1994) "Large sample estimation and hypothesis testing", *Handbook of Econometrics*, 4, 2111-2245.

19. Petrin A. (2002) "Quantifying the benefits of new products: The case of the minivan", *Journal of Political Economy*, 110, 705-729.

20. Seetharaman P. (2004) "Modelling multiple sources of state dependence in random utility models: A distributed lag approach", *Marketing Science*, 23(2), 263-271.

21. Shcherbakov O. (2009) "Measuring consumer switching costs in the television industry", Unpublished manuscript.

22. Sudhir K. and Yang N. (2014) "Exploiting the choice-consumption mismatch: A new approach to disentangle state dependence and heterogeneity", Working Paper.

23. Train K. (2009) "Discrete choice methods with simulation", *Cambridge University Press.*

24. Zimmermann A. and Heckelei T. (2012) "Structural change of European dairy farms, a cross-regional analysis", *Journal of Agricultural Economics*, 63(3), 576-603.

Appendix 2.A

Table 2.A.1 Overview of the Model

	Step 1	Step 2
Estimated on Markets	With Micro Data Only	All Markets
Estimation Method	Maximum Likelihood	GMM
Estimated Parameters	Product Character. Coeff. & Weights Loyalty Coeff. & Weights	Product Character. Coeff. & Weights -
Parameters Treated as Data	-	Loyalty Coeff. & Weights

Table 2.A.2: City B Product Characteristics

Variable	Mean	Std. Dev.	Min.	Max.	N
Share of Fat Free	0.399	0.173	0.157	0.682	424
Share of Light Products	0.386	0.312	0.008	0.911	424
Share of Regular Products	0.174	0.254	0	0.767	424
Share of Natural Labelled Products	0.175	0.327	0	1	424
Share of Greek Labelled Products	0.331	0.397	0	1	424
Share of Organic Products	0.169	0.320	0	1	424
Share of Original Labelled Products	0.061	0.091	0	0.262	424
Number of Flavors	46.672	40.99	2	130	424

2.A.3 Derivatives of the Micro Model

The derivative of the Likelihood function is written as:

$$L = \sum_h ln(\sum_r \sum_q w_r w_q \prod_t s_{hrqjt}^{y_{jt}})$$

$$\frac{\partial L}{\partial \theta_k} = \sum_h \frac{1}{\sum_r \sum_q w_r w_q \prod_t s_{hrqjt}^{y_{jt}}} \sum_r \sum_q w_r w_q \sum_t \frac{\partial s_{hrqjt}}{\partial \theta} \prod_{t' \neq t} s_{hrqjt'}$$

$$\frac{\partial s_{hrqjt}}{\partial \theta_k} = s_{jrqht}(\frac{\partial \delta_{jt}}{\partial \theta_k} + x_{jk} - \sum_{j'}(\frac{\partial \delta_{j't}}{\partial \theta_k} + x_{j'k}) s_{hrqj't})$$

$$\frac{\partial L}{\partial w_r} = \sum_h \frac{1}{\sum_{r'} \sum_q w_{r'} w_q \prod_t s_{hr'qjt}^{y_{jt}}} \sum_q w_q (\prod_t s_{hrqjt}^{y_{jt}} + w_r \sum_t \frac{\partial s_{hrqjt}}{\partial w_r} \prod_{t' \neq t} s_{hrqjt'})$$

$$\frac{\partial s_{hrqjt}}{\partial w_r} = s_{hrqjt}(\frac{\partial \delta_{jt}}{\partial w_r} - \sum_{j'} \frac{\partial \delta_{j't}}{\partial w_r} s_{hrqj't})$$

Where the derivatives from the mean utility follow from the implicit function theorem as described in Nevo (2000):

$$\frac{\partial\delta}{\partial\theta} = -1\frac{\partial S}{\partial\delta}^{-1} * \frac{\partial S}{\partial\theta}$$

$$\frac{\partial\hat{S}_t}{\partial\delta_t} = \sum_h \sum_r \sum_q w_r w_q (diag(s_{hrqt}) - s_{hrqt}s'_{hrqt})$$

$$\frac{\partial s_{jt}}{\partial\theta_k} = \sum_h \sum_r \sum_q w_r w_q s_{jrqht}(x_{jk} - \sum_{j'} x_{j'k}s_{j'rqht})$$

$$\frac{\partial\delta}{\partial w_r} = -\frac{\partial S}{\partial\delta}^{-1} * \frac{\partial S}{\partial w_r}$$

$$\frac{\partial s_{jt}}{\partial w_r} = \sum_h \sum_q w_q s_{jrqht}$$

Chapter 3

Technology Adoption, Vertical Restraints and Partial Foreclosure: Changing the Structure of an Industry

This chapter is co-authored with Alon Eizenberg and Michelle Sovinsky.

Acknowledgements: We thank Gus Aspará and Tobi Weber for excellent research assistance. For helpful discussions we thank Jan Boone, Brett Gordon, Saul Lach, and Konrad Stahl. We thank seminar participants at Edinburgh, Lausanne, Mannheim, Tel Aviv, Tilburg, and Zurich. We also thank Brett Gordon for sharing his data with us for earlier years.

3.1 Introduction

"There is perhaps no aspect of competition policy that is as controversial or has been as inconsistent over time and across jurisdictions as policy towards restraints between upstream firms and their downstream retailers" LaFontaine and Slade (2008)

Upstream manufacturers often impose exclusive dealing contracts on their retailers, which may result in foreclosure of a competing brand. On the other hand, there may be procompetitive effects of exclusive dealing. For example, exclusivity could enhance market performance by inducing a retailer to focus its promotional activities on the manufacturer's products and thereby enhance the provision of customers' service. It could also secure investments made by the manufacturer (such as quality assurance and advertising) by preventing the retailer from "free-riding" on these investments.[1] Due to its potential procompetitive effects, the *per se* illegality characterization of exclusive dealing was rejected in Standard Oil Co. v. United States (Standard Stations), 337 U.S. 293, 305 -06 (1949) and the US courts treat exclusive dealing under the "rule of reason" legality rule.[2] Motivated by these conflicting aspects of exclusive dealing contracts, a vast theoretical and empirical literature has sought to identify their impact given various market conditions.

One of the challenging tasks within this research agenda has been to empirically identify the foreclosure effect, i.e., the impact of vertical restraints imposed by one upstream firm on the sales of a competing upstream firm. An example of such work is Ater's (2010) documentation of the negative effect of exclusive dealings on upstream rivals' market shares in the fast-food industry. In general, empirical evidence on this issue is scarce. Our paper wishes to contribute to this literature while emphasizing an aspect that has not been addressed to date, to the best of our knowledge: the role played by vertical restraints in generating partial foreclosure in a dynamic setup.

[1] Marvel (1982) and Besanko and Perry (1993) provide more discussion of this issue.

[2] Potential pro- and anti-competitive effects of exclusive dealing and the history of its legal statues in the US are discussed in Areeda and Kaplow (1997) and Sullivan and Hovenkamp (2003). Exclusive dealing may violate the Clayton Act (Section 3) and the Sherman Act (Section 2). The rule of reason approach was reaffirmed in Tampa Elec. Co. v. Nashville Coal Co., 365 U.S. 320 (1961)

Consider an upstream industry that sells critical components to downstream customers. This upstream industry features a dominant incumbent firm, and a smaller entrant, whose stated goal is to challenge the dominant position of the incumbent. Suppose further that this upstream industry is characterized by rapid innovation, large sunk investments, and capacity constraints. Downstream customers, for their part, crucially depend on timely shipments from the upstream industry.

In this setup, exclusive dealing contracts with major downstream clients may allow the dominant firm to curb the downstream adoption of the rival's technology, effectively inhibiting its growth. Being excluded from selling to the major downstream customers, the smaller upstream supplier may struggle to expand its customer base and production volumes. This, in turn, may limit its ability to finance investments in Research and Development (R&D) and capacity, perpetuating its inferior position relative to the incumbent. This mechanism is likely to be taken into account by downstream customers: they may become skeptic about the smaller competitor's ability to deliver timely and reliable shipments, and these beliefs may make it even harder for the smaller upstream supplier to expand its sales. In this scenario, foreclosure takes a sophisticated and subtle form: rather than excluding the rival out of the market outright, the dominant firm uses exclusive dealing contracts in order to put in motion a cycle of forces that secure its dominant position.

This paper estimates dynamic panel models that document such patterns in the semiconductor industry. Specifically, we study competition in the microprocessor x86 market between Intel, traditionally controlling about 80% of the market, and its smaller rival AMD that controls much of the remaining 20% market share. Our analysis covers the years 2002-2010 and focuses on the role played by the "Intel Inside" program. Mostly throughout 2002-2005, this program funneled substantial payments from Intel to its downstream clients, i.e., PC manufacturers such as Dell, HP and IBM, in the form of rebates and advertising subsidies. As revealed in court cases and regulatory investigations, such payments were sometimes explicitly conditioned on the extent of the client's purchases from Intel's rival, AMD (see Lee, Péchy and Sovinsky 2013, hereafter LPS).

Intel's arrangement with HP, for example, specified that the share of HP's business line PCs using AMD's chips was not to exceed 5%, while manufacturers such as Dell and Toshiba refrained from using AMD's chips altogether.

The Intel Inside program illustrates the basic dilemma faced by regulators with respect to exclusive dealings. On the one hand, excluding AMD from downstream manufacturers' product lines clearly raises anticompetitive concerns. On the other hand, the program may have had procompetitive aspects: it provided downstream manufacturers with incentives to focus their entire production process around Intel's technology, possibly creating economies of scale with some of the efficiency gains being passed to consumers. Fierce downstream competition may have, indeed, allowed consumers to enjoy a substantial share of such efficiency gains, and rebates on CPU purchases may have similarly ended up benefitting consumers. This program may have also facilitated Intel's ability to capitalize on its investments, thus promoting innovation. Nonetheless, LPS find evidence that the advertising subsidies received by some of Intel's large customers (notably Dell) were *predatory* in the sense that their magnitude appeared inconsistent with pure profit maximization. The controversial nature of the program was manifested in a series of complaints and lawsuits filed by AMD with antitrust authorities and courts worldwide, leading to active investigations and lawsuits filed by regulators. This process gained strong traction in the years 2004-2005. Ultimately, the bulk of these legal proceedings were settled, with Intel agreeing to roll-back the controversial aspects of this program.

Ideally, empirical work should attempt to measure both the social benefits and the social costs of the Intel Inside program. Our goal in this paper, however, is more limited: we wish to identify the impact of this program on the evolution of competition in this market, and, specifically, on the process of downstream adoption of the AMD technology. While this is a less ambitious task than fully characterizing the effect of the program on welfare, it proves to be sufficiently challenging, due to the dynamic nature of technology adoption in this market.

Our empirical analysis combines several datasets. We document the downstream adoption of AMD's technology using market level data on PC brands' prices, characteristics and sales, indicating the share of PCs that has an AMD chip installed for each PC brand over time. We also collected data on the evolution of the upstream firms' technology and capacity. We draw on court cases and additional sources to construct variables that describe various aspects of the Intel Inside program, such as the volume of payments made to individual downstream clients, and specific restrictions dictated by Intel on their use of the AMD technology. Finally, we construct indices capturing the extent of litigation mounted by AMD and competition authorities worldwide in connection with the Intel Inside program. The joint variation of such variables over time allows us to identify the dynamic impact of technology, capacity, vertical restraints, and litigation, on the downstream adoption of AMD's technology. Sharp variation in some of these measures, such as Intel's decision to roll-back much of the controversial aspects of its Intel Inside program in 2005 (i.e., midway through the sample), is very helpful in separately identifying these effects.

We employ these data in estimating linear and nonlinear dynamic panel models in which the unit of observation is an individual downstream product line (e.g., HP's "Pavilion" desktop), and where the dependent variable is the share of this product line's sales that have an AMD chip installed. The remaining variables serve as explanatory variables. The following patterns emerge: the adoption of AMD's technology by a given downstream customer responds negatively to (i) the extent of payments via "Intel Inside" to the downstream customer itself, (ii) Intel Inside payments to Dell, (iii) specific restrictions on the extent of usage of AMD's technology imposed by the Intel Inside arrangement, and (iv) a measure of Intel's technology performance. We further find that the extent of adoption responds positively to (i) a measure of the AMD technology performance (ii) a measure of AMD's production capacity, and (iii) the extent of "anti-Intel" litigation.

We interpret these findings as reflecting the importance of dynamics in the technology adoption process. Institutional details suggest that this decision is inherently dynamic, for a number of reasons. First, as is typical with technology adoption, the extent of current adoption affects future costs of using the technology. For example, when a downstream PC maker sets up, or expands, a production line that is especially suited for AMD-based PCs, it may effectively reduce its cost of using the AMD technology in the next period. Other dynamic considerations have to do with the fact that the extent of current adoption of the AMD technology may affect both current and future benefits from the "Intel Inside" program. Ultimately, downstream customers need to weigh the potential benefits from adopting AMD's technology against its costs, and evaluation of both the benefits and the costs crucially depends on firm's expectations regarding several strategic variables: AMD's production capacity and technological progress contrasted with that of Intel, and the future viability of Intel's vertical restraints.

The fact that Intel's payments to Dell affect the adoption of AMD's technology by other downstream firms is noteworthy. Although the exact amounts paid to Dell were not common knowledge at the time they were made, PC makers may have had some information about them. If a downstream PC maker observes (albeit imperfectly) large payments from Intel to Dell, it may deduce that Intel is strongly determined to keep Dell as an exclusive customer. They may, therefore, use this information to update their beliefs regarding AMD's future market performance and its ability to serve as a viable supplier in the long run. The effect we document is, therefore, an interesting aspect of the dynamic and subtle effect of exclusive contracts on the downstream adoption of the rival technology.

The empirical evidence is consistent with a nuanced form of foreclosure: one not targeted at eliminating a rival, but rather at keeping this rival below a certain threshold of production, ensuring that it does not develop the capacity and size that would allow it to threaten the incumbent's dominant position.

The rest of the chapter is organized as follows. After a brief literature review, section 3.2 provides background regarding the industry. Section 3.3 describes the data sources we draw on. Section 3.5 describes the model. Section 3.6 explains our empirical strategy, i.e., the applicability of dynamic panel models to the question at hand. Section 3.7 provides our estimation results, and section 3.8 concludes with some discussion of limitations and future research.

3.2 Related Literature and Background

A vast theoretical literature provides motivation for pro as well as anticompetitive views of exclusive contracts and other vertical restraints. Exclusive dealing was initially considered an anticompetitive barrier to entry. However, the Chicago-critique maintained that an exclusive deal is not in the joint interest of the upstream and downstream firm, thereby dismissing the view that exclusive contracts could be used to exclude a rival and enhance market power.[3]

The post-Chicago theoretical literature, in contrast, has identified conditions under which socially harmful exclusive contracts may arise. These are situations in which an incumbent and a retailer may have a joint incentive to contract on exclusive dealing as a means of foreclosing entry. The main insight is that such contracts imply externalities on other players not accounted for in the Chicago-critique. Spector (2011) finds that if contracts are required to be simple enough, this strategy may induce inefficient exclusion even if the excluded firm is present at the contracting stage. Exclusive contracts may thus cause inefficient eviction, not only entry-deterrence, even though the former is less likely than the latter. Yehezkel (2008) considers the relationship between vertical contracts and product variety, and focuses on the case where a dominant manufacturer prohibits its retailer from selling products that compete with its own, and finds that quality choices by downstream retailers depend crucially on the extent to which the retailer is privately informed about consumers' average willingness to pay. Under full information, the retailer offers both quality varieties only if it is optimal to do so under vertical integration. However, when the retailer is privately informed about demand, it offers both varieties even if under vertical integration it is profitable to offer only the manufacturer's product. If the manufacturer can

[3] First, if offering a second brand increases the retailer's profit, then the manufacturer can charge the retailer higher franchise fees. Therefore, if a manufacturer finds it profitable to foreclose a competing brand then it has to be that this brand is not profitable to begin with. Second, if for whatever reason a manufacturer wishes to foreclose a competing brand, then the manufacturer can choose between imposing exclusive dealing on the retailer, or offering him quantity discounts to induce the retailer to choose not to carry the competing brand. Either way, the manufacturer needs to compensate the retailer for the foregone profits from offering the competing brand. Thus it is not clear why exclusive dealing is any better from the manufacturer's viewpoint than quantity discounts that are less open to antitrust scrutiny. Third, the fact that the retailer has the option to carry the competing brand will force the manufacturer to offer discounts that the retailer is likely to pass on, at least partially, to consumers. In that sense, the competitive pressure from the competing brand holds even in the presence of exclusive dealing. As Gilbert (2000) points out, the arguments made by the "Chicago School" parallels a more tolerant approach by US courts towards exclusive dealing.

impose exclusive dealing, it will do so and foreclose the low quality substitute even if under vertical integration it is profitable to offer both varieties.[4]

Our paper belongs to a small but growing empirical literature on exclusive dealing and vertical contracts (see Lafontaine and Slade 2008 for an overview). Asker (2005) examines the Chicago beer market and uses an indirect approach to measure the effect of exclusive dealing on entry. He finds that rivals do not have higher costs when they must compete with firms who sell under exclusive dealing agreements, in contrast to a raising-rivals'-cost view of exclusive dealings. Sass (2005) also studies the beer market and finds that exclusive dealing is more prevalent in smaller markets, in contrast to the predictions of foreclosure theory models. Nurski and Verboven (2013) estimate a structural model of demand with product and spatial differentiation and dealer exclusivity applied to the automobile market, asses the anti-competitive profit incentives for exclusive dealing, and subsequently evaluate the impact on consumers and welfare. They find that exclusive dealing in the European car market has served as a mild barrier to entry against Asian competitors, but with considerable consequences on consumers' domestic welfare because of reduced spatial coverage.

Our paper is closest to Ater (2012) who empirically quantifies the effect of exclusive dealing contracts on sales in the fast food industry. He finds that exclusive dealing reduces sales, and concludes that this is inconsistent with efficiencies, so that exclusive dealing must be used for anti-competitive reasons. Our paper contributes to this line of research by considering a technology market where an interesting and complex dynamic relationship arises between exclusive dealings, downstream technology adoption and upstream innovation and capacity investments.[5]

Finally, our paper joins a large literature on the PC and CPU industries. Several papers study the nature of innovation in the x86 microprocessor industry. Some examples that rely on static structural models include Song (2007), who uses the pure characteristics demand model to quantify the benefits from such innovation, and Eizenberg (2014), who studies the impact of CPU innovation on the variety of downstream PC configurations. Gordon (2009) uses a dynamic

[4] For related theoretical contributiuons, see for example Inderst and Shaffer (2007) and Lommerud, Straume and Stogard (2003).

[5] Additional contributions include Slade (2005) and Suzuki (2009).

demand model to study consumer replacement cycles, and Goettler and Gordon (2011) estimate a dynamic model in which innovation by Intel and AMD is endogenously determined, and use it to predict the impact of innovation from a hypothetical exclusion of AMD from the market. Our work differs from these papers by studying the dynamic adoption process of the AMD technology by downstream PC makers. Our empirical approach relies on dynamic panel methods rather than on structural modeling, which has both benefits and costs: while the lack of a dynamic structural model limits our ability to analyze out-of-sample scenarios, it allows us to avoid some of the strong assumptions required in the estimation of dynamic games. In particular, we are able to perform estimation on a dataset with rich and detailed product-level data without running into large state space concerns.

A closely related paper is LPS that uses a structural approach to estimate the marginal benefit to Intel from a dollar invested in the "Intel Inside" program. As this marginal benefit appear substantially lower than the marginal cost, LPS conclude that the subsidizing of advertising via this program has been predatory in nature, and was meant to exclude AMD. Our current paper completes the picture by documenting this exclusion effect, paying attention to the unique characteristics of this exclusion: in particular, its partial nature (i.e., it foreclosed AMD from selling to particular product lines, or limited its sales to others, but did not result in complete exclusion of AMD from the market), and its dynamic effects.

Our goal is to identify the effect of vertical restraints imposed by an upstream supplier (Intel) on downstream input adoption decisions. We, therefore, set up an econometric model that treats a PC brand-quarter combination as the unit of observation, and defines the extent of the usage of AMD chips as the dependent variable. Specifically, we use the fraction of CPUs used in a PC brand that are AMD-based. Defining the dependent variable in this way helps us link our analysis to the nature of Intel's vertical restraints, which often specified a cap on the extent of AMD chips used as a condition for eligibility to the Intel Inside program benefits. Intel's restrictions, for their part, are the key explanatory variable which effect we wish to capture.

The analysis differentiates between sales in the home and business segments, due to important differences among such segments. These differences span several dimensions: technology, nature of demand, and the strategic considerations of Intel. For instance, Intel was allegedly more concerned about AMD's growth in the business segment than in the home segment, and tailored its vertical restraints appropriately.

Several econometric challenges arise in this analysis. A major challenge to overcome is the dynamic nature of PC makers' decisions to adopt AMD chips. The dynamic nature of these adoption decisions stem from several reasons. First, beginning to use AMD chips (or expanding the extent of their adoption) requires a certain degree of investment by the PC maker. In particular, the PC maker must learn how to configure the hardware (e.g., the motherboard) to AMD's specifications, and set up (or expand) a production line that installs AMD CPUs in the produced PCs. Since AMD and Intel chips are not "pin-compatible," one cannot simply plug one of them instead of the other, and certain adjustments need to be made. Such an investment in the learning and internal organization needed to expand AMD purchases is not likely to be a static, period-by-period decision, but rather a cumulative process. Importantly, the infrastructure created in a given quarter (where in "infrastructure" we mean the accumulated know-how, experience and physical aspects of an AMD-based production line) is likely to be useful in future quarters as well, and to reduce the cost of adopting AMD in these future quarters. The adoption decisions are best viewed, then, as a dynamic process in which current investment decisions are taken given expectations regarding future market conditions, formed based on observing current values of key state variables, some of which are endogenous. The state-dependent nature of these decisions manifests itself in several fashions. In particular, they imply that the endogenous adoption decisions at time $t-1$ affects the adoption decision at time t, and this dynamic link must be addressed via dynamic panel methods.

Another issue involves the nature of the expectations formed by the PC maker at time t regarding market conditions in future periods. This nature dictates the set of state variables that we wish to include as explanatory variables of the current adoption decision. It is commonly believed that AMD's production capacity plays an important role in this regard. For a PC maker to be willing to place increased importance on AMD chips, it must believe that AMD would be able to meet high levels of demand. PC makers' emphasis on keeping thin inventories implies that this issue is crucial for the creation of a strategic dependence on AMD's chips.

Capacity constraints are, in general, an important aspect of the integrated circuit industry. During the sample period, AMD's market share hovered around 20%, and it strived to increase its production toward parity with Intel. Intel, however, enjoyed a substantial advantage in terms of production facilities (FABs). It is likely that AMD faced a "chicken-and-egg" dilemma: in order to increase its market share, it would have to expand its production capacity so as to convince downstream OEMs that they can rely on timely, large-scale shipments. Constructing chip production facilities, however, is extremely expensive, and with low sales, AMD may have lacked the financial capabilities to finance such investments. The market's expectations regarding AMD's ability to expand its capacity, therefore, may have played a crucial role in OEM's adoption decisions. We address this issue by including measures of AMD's capacity (and of its capacity relative to that of Intel). This allows important events, such as the opening of a new AMD production facility, to affect current decisions by downstream OEMs via its effect on their expectations regarding AMD's future growth potential.

Other key determinants of the downstream OEM's expectations involve the relative performance of AMD vs. Intel chips. If AMD is able to introduce a new chip with superior capabilities, it not only affects the PC maker's current benefit from installing this new chip (stemming from being able to offer a more attractive PC today), but also expectations regarding the future performance of AMD. Specifically, it increases the perceived probability that AMD would be able to increase its market shares, allowing it to expand its capacity and progress toward parity with Intel. We, therefore, include as state variables the relative performance of AMD chips vs. those of Intel (captured via performance benchmarks, as explained below). On a similar vein, we include as state variables affecting the current decision the AMD chip price, again relative to that of Intel. A low AMD price may capture not only current downstream benefits (in the form of reduced PC marginal costs), but can also capture improved production processes: CPU innovation often leads to CPUs that are both faster and cheaper to produce (via the reduction of the die size).

So, a lower AMD price may also indicate its ability to invest in new production technologies and improve its competitive standing, as well as its future survival and growth prospects.

Intel's vertical restraints also play an important role on the right-hand-side of our econometric specification. These restraints play two roles. First, they may directly affect the extent of adoption of AMD chips, since they imply that using more than a certain amount of AMD chips may result in a loss of substantial monetary transfers via the Intel Inside program. Second, they also affect expectations. If Intel aggressively increases payments via the program, it may signal to the market that it is determined to block AMD from expanding. If this signal is credible, it may discourage current adoption of the AMD chip, since downstream clients may be worried about creating a strategic dependence on a supplier that may not be able to deliver substantial volumes of quality chips in the future. We, therefore, include as explanatory variables not only the vertical restraints imposed on the PC firm in question, but also measures that capture the overall scope and magnitude of Intel's restraints and subsidies.

3.3 Data

We use data from several sources, containing information on PC and CPU sales and attributes, PC firms' advertising expenditures, measures of processor quality, AMD and Intel's production capacity, and the evolution of their technology. In addition, we use information on the nature of the vertical restraints Intel imposed on downstream PC firms, and the extent to which legal action has been taken against Intel related to their vertical restraints as noted in court cases.

Sales We use quarterly data on PC sales in the US home and business sectors available from Gartner Group. A unit of observation is defined as a combination of PC vendor (e.g., Dell), PC vendor brand (e.g., Inspiron), market segment (e.g., Home), CPU vendor (e.g., Intel), CPU family (e.g., Pentium 4) and quarter. We focus our analysis on the home and business segments.[6]

[6] All variables expressed in monetary terms were deflated using the quarterly consumer price index of the Bureau of Labor Statistics, basis set at the year 2000 USD.

We exclude Apple products as those exclusively used IBM's chips for much of the sample period, utilizing Intel's chips afterwards.[7] We obtain data on worldwide CPU sales from various web sources.[8]

Table 3.1: Descriptive Statistics (PC Brand Segment Quarter Level)

	Obs	Mean	Std. Dev.	Min	Max
Percentage AMD Sold	3508	0.13	0.25	0	1
PC Characteristics					
Price PC (1000$)	3508	1.02	0.45	0.24	3.69
Brand Advertising (1000$)	3508	1.01	3.10	0.00	29.14
Firm Advertising (1000$)	3508	8.16	16.46	0.00	87.80
CPU Characteristics					
AMD CPU benchmark/dollar, if non-zero	1234	7.31	4.54	1.75	22.65
Intel CPU benchmark/dollar, if non-zero	3377	4.65	3.86	0.00	30.00
Num. Quarters Brand/AMD family available, if non-zero	1234	5.42	3.20	1.00	19.00
Num. Quarters Brand/Intel family available, if non-zero	3377	6.36	3.80	0.00	30.00
Capacity Related Variables					
Free Cash (100M$)	3508	8.31	3.12	3.97	19.05
AMD Capacity Index	3508	8.08	3.59	3.00	13.00
Intel Capacity Index	3508	31.93	6.71	23.00	44.00
Exclusionary Restriction/Antitrust Related Variables					
Exclusionary Restriction Index	3508	0.92	2.02	0.00	6.00
Intel Payments to Dell (M$)	3508	147.38	154.15	13.37	603.05
Intel Payments to each PC Firm (M$)	3508	36.83	95.24	0.00	603.05
Number of Antitrust Cases Against Intel	3508	4.19	2.15	1.00	7.00
Number Pending Antitrust Cases Against Intel	3508	3.45	1.62	1.00	6.00

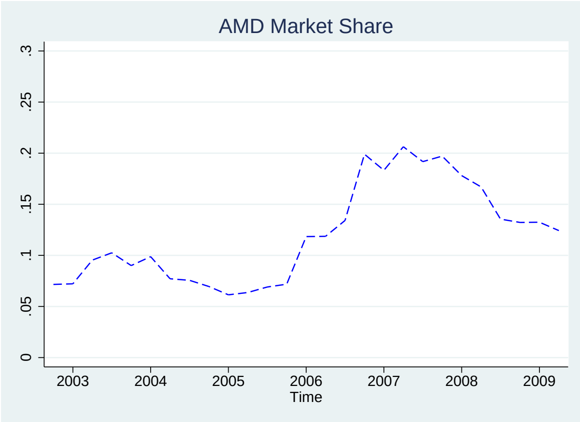
Table 3.1 shows descriptive statistics at the PC brand-segment-quarter level. The total number of observations is 3,508. As the table indicates, the rate of utilization of AMD's chips is on average 13%. The PC characteristics show significant variation as the average price of a brand ranges from $240 to $3,690. The average number of quarters a PC brand /CPU family pair is available on the market is one quarter higher for Intel than for AMD-based machines (conditional on the brand-segment offering at least one of the CPU manufacturer's products). Figure 3.1 shows the evolution of AMD's market (the remaining share garnered by Intel). AMD's share was below 10% until 2006, rising to 20% in the beginning of 2007, then declining again to 12.5% in 2009.

Figure 3.2 provides an overview of selected PC vendors' market shares and of AMD market share within each PC firm (remainder being Intel's market share). We show mean market shares

[7] The data did not record sales of servers prior to 2005 so we do not use information on servers (these are only 199 observations which is less than 4 percent of our data).

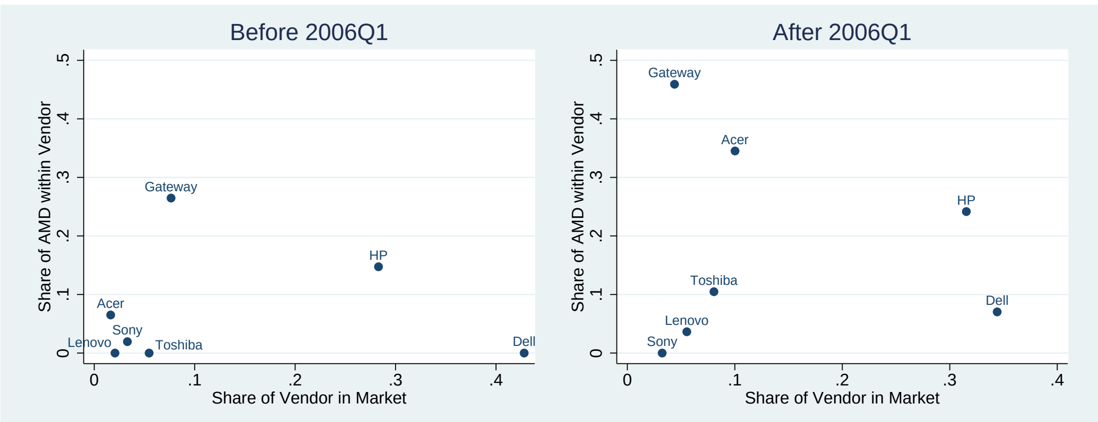
[8] Sources are: wikipedia.com, businesswire.com

Figure 3.1: AMD Market Share



over quarters on both home and non-home segments. We compute these on two different periods: before and after the first quarter of 2006 (a cut-off point after which most major restrictions were stopped by Intel). As can be seen, once the major restrictions were removed, the share of AMD CPUs within vendor increased at all vendors except Sony. The increase was close to 10 percentage points within the market leaders (HP and Dell), while it was larger than 20 percentage points within the smaller players (Gateway and Acer).

Figure 3.2: PC Vendor Specific Market Shares



Advertising PC advertisement data come from the Kantar Media Group. These data are important because many rebates given by Intel were a function of advertising done by PC firms. These consist of PC brand-specific ad expenditures (e.g. Acer Aspire) and PC firm level ad expenditures (e.g. Acer) where we match sales and advertising data across brands. Details about the definition of these variables and the matching to the Gartner data can be found in the appendix. Table 3.2 shows that expenditures for brand-specific ads averaged $1 million while firm-specific expenditures averaged $8.2 million (the observed maximum of $87 million was due to exceptionally large expenditures on TV advertisements). Table 3.2 provides an overview of the advertising variables. The brand advertising is summed at the vendor level on which the mean over time periods is computed. The firm level advertising is computed as the mean over time periods. As can be seen, PC firms have different advertisement expenditures. HP spends on average $6.9 million per quarter on brand advertising in the home segment and $37.9 million on firm level advertising. Dell, on the opposite, has very low advertisement spendings ($0.06 and $3.04 million respectively). The ratio of brand level versus firm level advertising differs across firms. Some of the firms spend more at the firm level than at the brand level (Dell, Gateway, HP, Lenovo), some spend more at the brand level (Sony, Toshiba). These variations allow us to identify the advertisement coefficients in the regressions.

Table 3.2: Quarterly Average Advertisement Spendings per PC Vendor, in $ mios

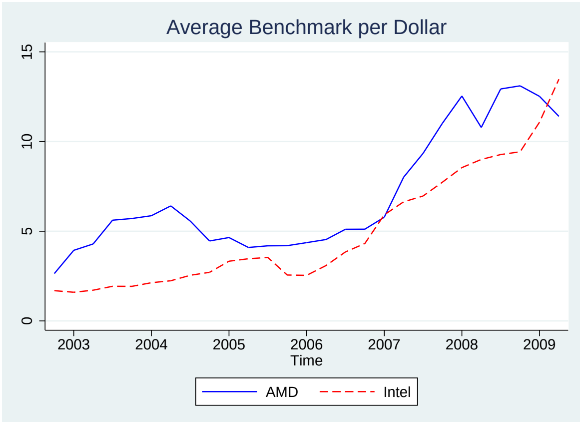
PC Vendor	Total Brand Level Advertising		Firm Level Advertising
	Home Segment	Non-home segment	
Acer	0.48	0.66	0.42
Dell	0.06	0.11	3.04
Gateway	0.12	0.15	4.62
HP	6.92	7.22	37.90
Lenovo	8.10	8.81	19.88
Sony	3.00	3.32	0.85
Toshiba	1.63	2.10	0.78

CPU Quality We obtain data on the quality of the CPU from Passmark CPU Mark publications. These contain a continuous quality measure ("benchmark") for each CPU model. We use this together with CPU prices, which we gathered from published list prices or obtained from Instat, to construct a benchmark per dollar spent index.[9] The descriptive statistics proposed in Table 3.1 are computed on the non-zero observations only in order to allow a meaningful comparison of the two variables since a PC brand without any AMD CPUs will always have a zero value for its AMD quality measure. We see that AMD CPUs have on average almost 35% better benchmark per dollar than Intel CPUs. Moreover, Figure 3.3 shows the evolution over time of

[9] A detailed description of the data collection and aggregation can be found in the appendix.

the sales weighted average of the benchmark per dollar measure for both firms. As can be seen, AMD's quality measure was consistently better than Intel's throughout the sample, both due to lower AMD prices and higher AMD benchmark values.

Figure 3.3: CPU Quality per Dollar



CPU Production Technology and Capacity To account for production technology/total capacity of the CPU manufacturers, we use information from Intel and AMD annual reports. For each year, these report: the number of fabrication units available for making microprocessors (FABs), the silicon wafer size of each FAB (the larger the wafer, the more CPUs can be printed simultaneously), and the precision in nanometers available at each FAB for "printing" the integrated circuits (the smaller the precision, the more CPUs can be printed, moreover the CPU power efficiency is improved). We construct a measure of output/quality of the technology process of Intel (respectively AMD), the capacity index, as the sum over all FABs of (ordered) wafer size and (ordered) IC process. We also construct a measure of AMD's "cash flow available" for investing at the beginning of each quarter (from SEC quarterly reports).[10] Table 3.1 reveals that the Free Cash available for AMD in each quarter is on average \$831 million and that the mean capacity index of Intel is on average 4 times larger than AMD's. Table 3.3 sheds light on this difference in capacity indices: it shows the evolution of the CPU manufacturers' production technology and capacity over time. AMD usually lags behind Intel regarding the IC process and the wafer size.

[10] The quarterly filings were accessed on September 18, 2014 from http://ir.amd.com/

The variation in Intel's number of FABs is due to technological upgrades and relocations. As can be seen, our capacity index captures these variations, and demonstrates Intel's established advantage in production technology and capacity.

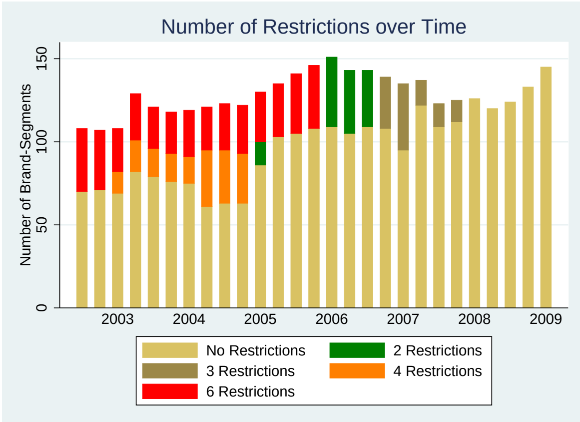
Table 3.3: Evolution of Capacity Variables

Year	Number of Fabs		Mean ICP in nm		Mean wafer in mm		Capacity index	
	AMD	Intel	AMD	Intel	AMD	Intel	AMD	Intel
2002	1	10	130	150	200	220	3	28
2003	1	7	130	130	200	229	3	23
2004	1	7	130	113	200	243	3	27
2005	2	6	90	78	250	300	9	33
2006	2	5	90	75	250	300	9	28
2007	2	5	78	57	250	300	10	32
2008	2	7	65	60	300	300	12	44
2009	2	6	55	50	300	300	13	41

Notes
1. ICP stands for Integrated Circuit Process
2. The Fab capacity index is computed by ranking IC process (largest to smallest) and wafer size (smallest to largest) , then summing these points over all fabs

Exclusive Restrictions We examined case files from the Department of Justice and the European Commission that listed the types of exclusive restrictions Intel imposed on downstream PC firms. From these files we could categorize the type of restrictions imposed, which include a cap on amount sold of AMD or restrictions on sales in specific segments/product lines, restrictions on distribution channels used for AMD products, provision of rebates in exchange for selling certain amounts of Intel-based machines, limitations on the marketing PC firms could undertake for AMD-based products, guarantees of preferred supply of Intel CPUs, and restrictions imposed on bidding on contracts using AMD-based products. These documents also detail threats made by Intel to certain PC firms to remove funding or channel funding to rivals or other retaliation as a punishment for selling more than the specified amount of AMD chip. Based on these restrictions we form two indices - one is the sum of the number of restrictions imposed, and one is the sum of the number of more extreme restrictions imposed. The latter includes: excessive rebates, demands to exclude AMD from certain product lines completely, threats, or a promise to increase Intel market shares provided to the PC firm. Table 3.1 reveals that, on average across brand-segment-quarters, 1.03 restrictions were in place while the maximum was 6. Figure 3.4 shows the number of brand segments which were affected by the exclusionary restrictions index in each quarter. As can be seen, most of the restrictions took place before 2007. The variation reflected in this figure allows the identification of the coefficient of the index in our model.

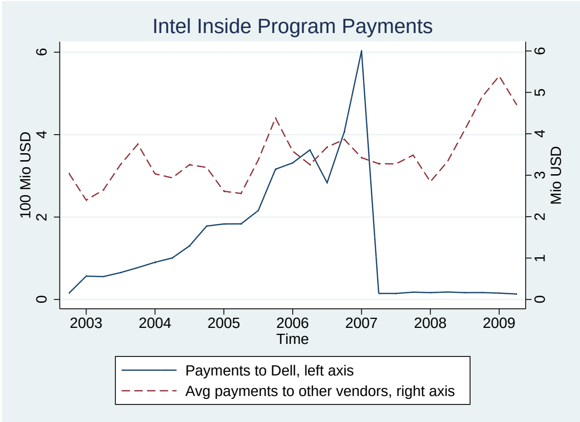
Figure 3.4: Evolution of Restrictions Imposed by Intel



We also observe rebates offered by Intel to PC vendors via the Intel Inside program. In the case of Dell, the amounts of the exact payments are available from the court decision SEC vs. Dell Inc. July 22 2010, covering the years 2003 until 2007. After 2007 we refer to the official advertisement of the Intel Inside program on Intel's webpage. The program stated that 3% of the CPU costs of the PC vendor would be offered by Intel to finance ads for PC models equipped with Intel CPUs.[11] Hence, for the period after 2007, the Intel Inside payments are computed as 3% of Dell's CPU costs (computed using Gartner sales data and the price dataset we described previously). An identical approach, using a 3% rebate, is used to compute the Intel Inside payments received by other PC firms. The variable is defined at the firm level and summed over all brands and segments. Figure 3.5 proposes an overview of these Intel Inside payments. The payments are sorted into two groups: the payments to Dell on the left axis and the average payments to the other PC vendors on the right axis. First, we note that the payments to Dell (over the 2003-2007 period) are close to 100 times the payments the company was supposed to receive based on the advertised 3% rebate. Examining payments to other PC vendors (computed based on the 3% rebate), it appears that inclusion in the Intel Inside program conveyed non-negligible benefits: these payments averaged between \$2.3 and \$5.5 million per quarter per firm.

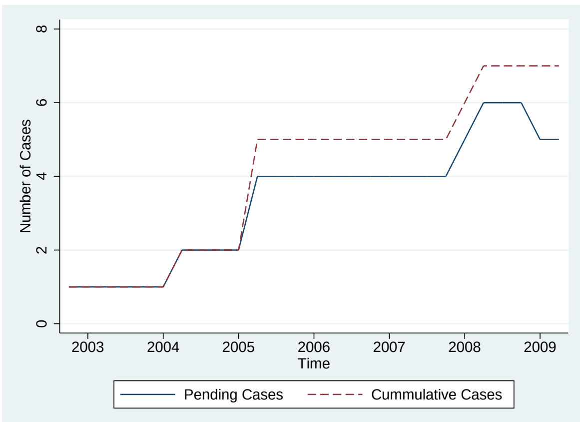
[11] For further details about the program, see for example Lee, Pchy and Sovinsky (2013).

Figure 3.5: Evolution of Intel's Payments to PC Vendors



Legal Action We compiled information on investigations or legal action taken against Intel relating to its practices against AMD. We construct two variables based on lawsuits filed by the Federal Trade Commission, European Commission, Korean Fair Trade Commission, Japanese Fair Trade Commission, State of New York, and found in the shareholder reports of both Intel and AMD. We include measures of the cumulative number of antitrust cases/investigations brought against Intel as of 2001 and the number of pending antitrust cases/investigations in process against Intel. These variables may capture the extent to which market participants revise their beliefs regarding the future viability of Intel's restraints. Table 3.1 shows that on average across all quarters there were 3.45 pending cases, and 4.19 cumulative cases. Figure 3.6 provides the evolution over time of these two measures. The majority of the cases were launched by 2006, and the first settlement occurred in the first quarter of 2005.

Figure 3.6: Legal Action Against Intel



3.4 Intuition

In this section, we propose a simple theory model to expose the dynamic factors affecting a PC firm's decision on CPU adoption. For simplicity, the behaviour of CPU manufacturers is not modelled and treated as exogenous. Time is discrete. The one period profit function of the PC firm can be written as:

$$\begin{aligned}\Pi(W_{it}, x_{it}|c_{t-1}, m_{it}, r_{it}) =& \text{Revenue}(W_{it}, x_{it}, m_{it}) - \text{costofPC}(x_{it}) - \text{costofCPU}(W_{it}, c_{t-1}, m_{it}) \\ & - \text{FirstAdoptionFC}(W_{it}) + \text{IntelInside}(W_{it}, r_{it}).\end{aligned}$$

The PC firm can influence its period profit through two choice variables: the share of AMD CPUs W_{it} and the PC characteristics x_{it}. Moreover, its profit is also affected by the values of the state variables: the lagged production capacity of the CPU firms c_{t-1} (a vector containing measures for both Intel and AMD), the characteristics of the CPUs m_{it} (a vector containing measures for both Intel and AMD) and the restrictions imposed by Intel r_{it}. The profit function is broken down to the following elements: the revenue generated by the PCs sold , the production cost of the PCs, the purchase cost of the CPUs, the restrictions imposed by Intel through the "Intel Inside" program and finally the First Adoption Fixed Costs incurred in case the PC firm is using AMD CPUs for the first time.

The state variables' values change over time as CPU manufacturers bring product improvements. The laws of motion are defined as:

$$\begin{cases} c_t = c_{t-1} + f^c(W_{it}) \\ m_{it+1} = m_{it} + f^m(W_{it}) \\ r_{it+1} = r_{it} + f^r(W_{it}) \end{cases}$$

The three functions $f(W_{it})$ take into account the reaction of the CPU manufacturers to the previous period purchase decision made by the PC firm. Let us take as an example the case where the PC firm purchases more AMD in period t (larger W_{it}). The two CPU manufacturers will react in in $t+1$ in different ways. For AMD, this translates in extra profit, which can be invested in production capacity (c_t) and/or in product quality (m_{it}). For Intel, it suggests that their clients are not following their "Intel Inside" agreements and therefore Intel will likely react by decreasing benefits (increasing restrictions) (r_{it}) on the PC firm.

PC firms are forward looking and therefore aware that the decisions (W_{it}) they are taking today will affect the state variables they will be facing tomorrow. However, they have imperfect knowledge of the functions $f(W_{it})$. Therefore, they consider $f(W_{it})$ as stochastic processes and take their decisions based on $E(f(W_{it}))$. PC firms solve an infinite horizon dynamic programming problem. Given the period utility function, the state variables, and a discount factor β the optimal strategy of a PC firm is the solution to the following Bellman equation:

$$V(c_{t-1}, m_{it}, r_{it}) = Max_{W,x}\Pi(W_{it}, x_{it}|c_{t-1}, m_{it}, r_{it}) + \beta EV(c_t, m_{it+1}, r_{it+1}|c_{t-1}, m_{it}, r_{it})$$

where the value function $V(c_t, m_{it}, r_{it})$ is the solution to the equation. The solution to this dynamic problem are two policy functions setting the optimal share of AMD CPUs $W(c_t, m_{it}, r_{it})$ and the optimal PC charactersitics $x(c_t, m_{it}, r_{it})$. This model offers a synthesized overview of the key aspects of the behaviour of the PC firms. However, we do not attempt to estimate it directly as the data needs are very demanding. Instead, we propose a reduced form approach with a unique equation to be estimated, where the dynamics will be accounted for using the lag of the share of AMD within a PC vendor, W_{it} .

3.5 Econometric Model

We take as a unit of observation (i) the PC brand-quarter-segment (e.g., Acer's Home Market Aspire), where the extent of usage of AMD chips at time t is the dependent variable, W_{it}. In other words, it is the fraction of AMD-based CPUs used by a PC firm in each of its brands. This links our analysis to the nature of Intel's vertical restraints, which often specified a cap on the amount or percentage of AMD chips used by the PC firm as a condition for eligibility of the Intel Inside program benefits. We consider the home and business markets separately as some of Intel's vertical restraints focused explicitly on the business segment, while others were more general. Specifically, our model of a PC firm's choice regarding how much AMD to sell is given by

$$W_{it} = \alpha W_{it-1} + \beta x_{it} + \lambda c_{t-1} + \eta m_{it} + \delta r_{it} + \gamma l_{it} + \mu_i + \varepsilon_{it.} \tag{3.1}$$

Following the above discussion, W_{it-1}, the lagged percentage of segment-brand i sold with an AMD processor, serves as an explanatory variable to help us capture the dynamic link between current and future decisions. The μ_i represents brand fixed effects and ε_{it} is an idiosyncratic i.i.d. error term.

Time varying observed characteristics of PC brand i are captured by x_{it}.[12] These include the price of brand i, brand-level advertising, and firm-level advertising. Given that often more than one product is available within the brand, when the the component of x_{it} varies within i we compute the weighted average within the brand.

Variables relating to (lagged) CPU manufacturing capacity constraints are included in c_{t-1}. To capture upstream production technology investments (along with their effect on expectations), we include a capacity index for AMD and for Intel, which is a function of the number of FABs, wafer size and IC process (as described in section 3.3), as well as the amount of free cash available

[12] We do not include product characteristics that are not time varying within brands (such as the platform/form factor) because these characteristics are encompassed by the brand fixed effect term which will be eliminated in the econometric estimation methodology.

to AMD for investment, which captures the fact that AMD was more cash constrained than Intel.[13] One could argue that leads of capacity should be included, since the knowledge that AMD is about to expand its capacity may also affect decisions. On the other hand, one may argue that announcing that the new IC process would become "operational" on a given date, and announcing that it is operational on a given date, are very different pieces of information, which may justify using lagged values but not leads.

CPU related variables are given in m_{it}. These include the extent of technological progress of AMD as measured by the AMD CPU benchmark per CPU dollar, Intel CPU benchmark per dollar, the number of quarters the segment-brand-CPU family combination for those families sold by AMD has been available, and the number of quarters the segment-brand-CPU family combination for those families sold by Intel has been available[14].

The vector r_{it} captures upstream restrictions and can be conceptually decomposed into two parts: the part related to exclusive restrictions on PC sales imposed by Intel and the part related to the Intel Inside program, where the latter may have been used to enforce exclusive deals. Identifying the impact of the upstream restrictions on product line choice is the goal of our analysis. Specifically, the r_{it} contain total payments Intel made to Dell during the period 2002-2005 and 3% of CPU sales post 2005, total Intel inside payments to all firms, and an index that captures the intensity of exclusionary restrictions imposed by Intel (as discussed in section 3.3).

Finally, the vector l_{it} captures legal issues that may impact whether AMD is sold. These include the cumulative number of antitrust cases/investigations brought against Intel as of 2001 or the number of pending antitrust cases/investigations in process against Intel.

[13] Measures of capacity refer to worldwide capacity for production, while our variable of interest relates to sales in the US, which mitigates endogeneity concerns.

[14] We do not have information prior to 2001 so the number of quarters available is counted starting from the first quarter in 2001.

3.6 Estimation Methodology

Our goal is to identify the causal effect of upstream vertical restraints on downstream product line choices (i.e., firm's input decisions). Our estimation methodology takes advantage of the dynamic panel aspect of our data. This approach is valuable for a number of reasons. First, a dynamic panel can help us address unobserved firm heterogeneity in the decision to adopt AMD's chips. Unobserved heterogeneity may arise if some firms are fundamentally better-suited to gain from using AMD's chips. For example, this may stem from firms' natural positioning as "value PC" makers, which motivates them to offer PCs with cheaper CPUs installed (noting that AMDs prices are typically lower than those of Intel). In addition, some firms may have more flexible production processes than others (for instance, in the sense that they do not enjoy huge economies of scale from using only one type of chip).

One thing we need to address is that the use of vertical restraints may be endogenous to the outcome variable. The panel data are crucial to our identification of the causality of vertical restraints on input decisions. For example, endogeneity could arise when Intel has to incur substantial investments in order to produce their CPU chips, and, hence, they find it profitable to implement an exclusive deal. However, these investments that prompt vertical restraints, may separately influence PC firms costs and therefore our dependent variable. In the absence of an instrument (that is correlated with the use of ED but not with quantity) cross-sectional data would only allow us to uncover correlations between vertical restraints and the dependent variable. The panel aspect of our data allows us to overcome the endogeneity of ED by using a fixed-effect type of estimator. This removes time-invariant unobserved firm heterogeneity that is the cause of the endogeneity problem. Thus, the effect of the vertical restraint on the outcome variable, as captured in δ is causal and is identified by time-series variation. Note that we need time-series variation in the use of exclusive constraints so we need a panel of data that comes before and after the vertical restraints where implemented.

The relationship described in equation (3.1) is estimated under various assumptions and with a variety of specifications. We begin by considering the estimation of the relationship in equation (3.1) using a linear specification followed by a non-linear specification (Tobit). The Tobit enables us to consider explicitly the issue of zero-adoption observations. Both linear and non-linear approaches address endogeneity of the past-adoption decision W_{it-1} and endogeneity of the exclusive deals (r_{it}). Furthermore, the approach also accounts for the issue that components of the x_{it} and f_{it} may also be endogenous as they relate to product line choices.

3.6.1 Linear specification

Our linear strategy treats μ_i as a fixed effect to be differenced out following Arellano and Bond (1991, AB). Differencing out the fixed effects, we obtain a first-differences regression of the form:

$$\Delta W_{it} = \alpha_L \Delta W_{it-1} + \beta_L \Delta x_{it} + \delta_L \Delta r_{it} + \lambda_L \Delta c_{t-1} + \eta_L \Delta m_{it} + \Delta \varepsilon_{it}. \tag{3.2}$$

Following Wooldridge (2002), first we assume that the variables represented by $y_{it} \equiv \{x_{it}, r_{it}, c_t, m_{it}\}$ are strictly exogenous. That is, we assume ε_{is} is independent of y_{it} for any t, s conditional on μ_i. Strict exogeneity implies Δy_{it} is exogenous and hence Δy_{it} can be it's own instrument. However, Δy_{it} may be a weak instrument.[15] Another solution to potentially weak instruments is to assume that ε_{it} is independent over time, then the (two period) lag of the endogenous variable(s) is a valid instrument.[16]

However some of the regressors, even if independent of current disturbances, may be influenced by past ones. These regressors are then not strictly exogenous but rather exhibit sequential exogeneity where $E(\varepsilon_{it} \mid y_{is}, \mu_i) = 0$ for $s \leq t$. Relaxing the assumption of strict exogeneity implies Δy_{it} is endogenous. In this case we can use W_{it-2} and y_{it-1} as IV in the first-differenced equation.[17]

[15] We can test the null hypothesis that the overidentifying restrictions hold. This is a Wald statistic that the instruments are valid, which is valid under heteroskedasticity and clustering. The critical value is $\chi^2(l)$, where l is the degree of overidentification.

[16] In this case it is important to test the null hypothesis that ε_{it} is independent over time. We implement an autocovariance test of the null hypothesis of no autocorrelation in the idiosyncratic error term.

[17] We can test for weak instruments using the standard first stage regression results: if y_{it-1} are not weak instruments then they should affect W_{it-1} conditional on y_{it}. Again, we can test for serial correlation in the errors.

Note that estimators using too many lags of explanatory variables (relative to the number of observations) as instruments are known to have poor finite sample properties (Arellano and Bover (1995), and Blundell and Bond (1998)). In practice best not to use lags back to $t = 1$. We follow this approach and consider specifications using only lags of three periods. Furthermore, autocorrelation tests indicated that there is serial correlation in our error structure for two period lags, hence we use W_{it-3} and y_{it-2} as IV.

3.6.2 Nonlinear specification

The linear specifications above allowed us to control for both state dependence, and for individual heterogeneity, in explaining the rate of adoption of AMD's chips. These specifications, however, did not address the "corner solution" issue: many product lines, at different times, chose not to use AMD chips at all.

To address this issue, we follow Wooldridge (2002) in specifying a dynamic nonlinear model that builds on Chamberlain (1984). This allows us to include random effects that capture time-constant heterogeneity, as well as a lagged dependent variable capturing state dependence. However, this framework does not allow us to relax a strict exogeneity assumption. In other words, both the linear and the nonlinear models have their specific strengths and weaknesses. By considering both approaches our goal is to provide a more complete picture of the issues at hand.

This specification treats the dependent variable W_{it} similarly as in the linear specifications: it is a continuous measure with a mass point at zero given by

$$W_{it} = max\big(0, \alpha_{NL} W_{it-1} + \beta_{NL} x_{it} + \delta_{NL} r_{it} + \lambda_{NL} c_{t-1} + \eta_{NL} m_{it} + \mu_i + u_{it}\big) \tag{3.3}$$

$$u_{it} | (\overline{y}_i, W_{i,t-1}, ...W_{i0}, \mu_i) \sim N(0, \sigma_u^2), \tag{3.4}$$

where, as in the linear specification, we denote by y_{it} the collection of all the explanatory variables in all time periods. The mean (over time) of these variables is $\overline{y}_i$.

One issue concerns the initial value of W_{i0}. A possibility is to treat it as nonrandom, which would imply that μ_i and W_{i0} are independent. However, this may not necessarily be the case, so we specify the density of the fixed effect *conditional on the initial condition.* That is, we specify the fixed effects as

$$\mu_i = \psi + \xi_0 W_{i0} + \overline{y}_i \xi + a_i,\ a_i| \left(W_{i0}, \overline{y}_i\right) \sim N(0, \sigma_a^2) \tag{3.5}$$

As discussed in Wooldridge (2002), the fixed effects can then be integrated to yield the likelihood function of the random effects Tobit model with time-t, observation-i explanatory variables: $(y_{it}, W_{i,t-1}, W_{i0}, \overline{y}_i)$. That is, $\overline{y}_i$ and w_{i0} are controlled for in each time period. We will use this likelihood function to obtain estimates of the parameters $(\alpha_{NL}, \beta_{NL}, \delta_{NL}, \lambda_{NL}, \eta_{NL}, \psi, \xi_0, \xi, \sigma_a^2)$.

3.7 Results and Implications

Our results, reported in Tables 3.4 and 3.5, reveal that both the AB and Tobit estimations lead to qualitatively similar results. The lag of the dependent variable is positive and significant in both analyses. The PC price coefficient is significant, negative and stable across all specifications and for both approaches. The advertisement variables are not significant except in the case of brand advertising in the AB estimation where it has a negative effect on the share of AMD set by the PC firm. The CPU quality measures of AMD and Intel are stable and significant in all specifications and both types of regressions, moreover they have the expected effect on the adoption of AMD CPUs: positive for AMD CPU quality and negative for Intel CPU quality. The age of PC brand-CPU family pairs is insignificant across all specifications in AB but significant and stable in the Tobit regressions. The lag of free cash is insignificant under both approaches. The capacity indices are significant and have the expected signs in both tables (except for the

Intel capacity index, which is insignificant in the AB specifications): the AMD capacity index has a positive effect on the dependent variable, while the Intel capacity index has a negative effect.

We now discuss the key variables of interest: variables relating to the Intel Inside program. The adoption of AMD's technology by a PC firm responds negatively to the extent of Intel Inside payments to the downstream customer itself in both tables, and it is significant in the AB estimation. The same is true for the Intel Inside payments to Dell, whose coefficients are negative in both tables and significant in the case of the AB estimation. The specific restrictions on the extent of usage of AMD's technology imposed by the Intel Inside arrangement have a negative and significant effect in both tables, however with a lower significance level in the Tobit estimation (10%). We note that in specifications 10,11 and 12, when all variables related to the Intel Inside case are active simultaneously, the significance is lost (except for the Intel Payments to PC firms in the Tobit regression). Finally, the "anti-Intel Inside" litigation variables have a significant effect in the case of the Tobit regressions only. As anticipated, the effect is positive: the pressure of antitrust authorities on Intel encourages PC firms to increase their share of AMD equipped PCs.

These findings confirm our hypotheses laid down in the introduction. The restrictions imposed by Intel on the downstream players affect their supplier choices, and thus limit the expansion of the rival AMD on the market. Moreover, the number of pending and antitrust cases suggest that PC firms have dynamic considerations when choosing a CPU supplier. A stronger litigation against Intel today suggests better opportunities for AMD tomorrow, thus increasing interest of the PC firm to collaborate with AMD today.

Table 3.4: Arellano-Bond Regressions of Percentage of AMD Sold

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lagged Percentage AMD Sold	0.9345***	0.9349***	0.9348***	0.9342***	0.9345***	0.9345***	0.9342***	0.9345***	0.9345***	0.9342***	0.9345***	0.9344***
	(0.0195)	(0.0194)	(0.0195)	(0.0195)	(0.0194)	(0.0195)	(0.0195)	(0.0194)	(0.0195)	(0.0198)	(0.0197)	(0.0198)
PC Characteristics												
Price PC (100$)	-0.0015***	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0015***	-0.0016***	-0.0016***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Brand Advertising (10M$)	-0.0058***	-0.0058***	-0.0058***	-0.0063***	-0.0063***	-0.0063***	-0.0063***	-0.0063***	-0.0063***	-0.0068***	-0.0068***	-0.0068***
	(0.0018)	(0.0018)	(0.0018)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
Firm Advertising (10M$)	-0.0009	-0.0009	-0.0009	-0.0012*	-0.0012*	-0.0012*	-0.0012*	-0.0012*	-0.0012*	-0.0018*	-0.0018*	-0.0018*
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0010)	(0.0010)	(0.0010)
CPU Characteristics												
AMD CPU benchmark/dollar (10,000$)	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***	0.0033***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Intel CPU benchmark/dollar (10,000$)	-0.0070***	-0.0072***	-0.0071***	-0.0071***	-0.0072***	-0.0072***	-0.0071***	-0.0072***	-0.0072***	-0.0071***	-0.0072***	-0.0071***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Number Quarters Brand/AMD family available	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Number Quarters Brand/Intel family available	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Capacity Related Variables												
Lagged Free Cash (1000M$)	0.0085	0.0077	0.0074	0.0084	0.0076	0.0074	0.0084	0.0076	0.0074	0.0083	0.0075	0.0072
	(0.0095)	(0.0096)	(0.0096)	(0.0095)	(0.0096)	(0.0096)	(0.0095)	(0.0096)	(0.0096)	(0.0095)	(0.0095)	(0.0096)
AMD Capacity Index	0.0022**	0.0026***	0.0028***	0.0024***	0.0028***	0.0029***	0.0024***	0.0028***	0.0029***	0.0026***	0.0029***	0.0031***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0010)	(0.0010)	(0.0010)
Intel Capacity Index	-0.0000	0.0000	-0.0000	-0.0000	0.0001	-0.0000	-0.0000	0.0001	-0.0000	-0.0000	0.0000	-0.0000
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0005)
Exclusionary Restriction/Antitrust Related Variables												
Exclusionary Restriction Index	-0.0006**	-0.0006**	-0.0006**							0.0449	0.0458	0.0464
	(0.0003)	(0.0003)	(0.0003)							(0.1250)	(0.1250)	(0.1251)
Intel Payments to Dell (100M$)				-0.0024***	-0.0023***	-0.0023***				-0.0488	-0.0496	-0.0502
				(0.0006)	(0.0006)	(0.0006)				(0.1248)	(0.1248)	(0.1249)
Intel Payments to PC Firms (100M$)							-0.0023***	-0.0023***	-0.0023***	0.0008	0.0008	0.0008
							(0.0006)	(0.0006)	(0.0006)	(0.0008)	(0.0008)	(0.0008)
Number Pending Antitrust Cases Against Intel		-0.0023			-0.0022			-0.0022			-0.0021	
		(0.0019)			(0.0019)			(0.0019)			(0.0019)	
Number of Antitrust Cases Against Intel			-0.0018			-0.0017			-0.0017			-0.0017
			(0.0015)			(0.0015)			(0.0015)			(0.0015)

Notes: Robust Clustered Standard Errors in Parentheses. Each regression contains a constant and a quarterly trend. *** denotes significance at the 1% level, ** at the 5% level; and * at the 10% level. Each regression contains 3106 observations.

Table 3.5: Tobit Regressions of Percentage of AMD Sold

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Lagged Percentage AMD Sold	0.877***	0.875***	0.874***	0.868***	0.866***	0.684***	0.868***	0.866***	0.865***	0.860***	0.858***	0.858***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.029)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
PC Characteristics												
Price PC (100$)	-0.023***	-0.022***	-0.022***	-0.024***	-0.022***	-0.033***	-0.024***	-0.022***	-0.022***	-0.024***	-0.022***	-0.023***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Brand Advertising (10M$)	-0.027	-0.027	-0.026	-0.033	-0.033	-0.034	-0.033	-0.033	-0.033	-0.037	-0.037	-0.036
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Firm Advertising (10M$)	0.009**	0.010**	0.010**	0.005	0.005	0.006	0.004	0.005	0.005	0.002	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
CPU Characteristics												
AMD CPU benchmark/dollar (10,000$)	0.017***	0.018***	0.018***	0.017***	0.018***	0.017***	0.017***	0.018***	0.018***	0.017***	0.018***	0.018***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Intel CPU benchmark/dollar (10,000$)	-0.017***	-0.017***	-0.017***	-0.017***	-0.017***	-0.017***	-0.017***	-0.017***	-0.017***	-0.018***	-0.017***	-0.017***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number Quarters Brand/AMD family available	0.010***	0.010***	0.010***	0.009***	0.010***	0.013***	0.009***	0.010***	0.010***	0.010***	0.010***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number Quarters Brand/Intel family available	-0.007***	-0.008***	-0.008***	-0.007***	-0.008***	-0.009***	-0.007***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Capacity Related Variables												
Lagged Free Cash (1000M$)	0.024	0.030	0.033	0.017	0.023	0.026	0.015	0.022	0.024	0.010	0.016	0.018
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
AMD Capacity Index	0.017***	0.016***	0.013***	0.018***	0.016***	0.013***	0.018***	0.016***	0.014***	0.018***	0.016***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Intel Capacity Index	-0.003***	-0.004***	-0.003***	-0.003**	-0.003***	-0.003***	-0.002**	-0.003***	-0.003***	-0.002**	-0.003***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Exclusionary Restriction/Antitrust Related Variables												
Exclusionary Restriction Index	-0.008*	-0.007*	-0.008*							0.001	0.002	0.001
	(0.004)	(0.004)	(0.004)							(0.006)	(0.006)	(0.006)
Intel Payments to Dell (100M$)				-0.222	-0.257	-0.765*				0.725	0.669	0.734
				(0.151)	(0.189)	(0.420)				(0.519)	(0.528)	(0.522)
Intel Payments to PC Firms (100M$)							-0.217	-0.245	-0.224	-0.966*	-0.939*	-0.979**
							(0.135)	(0.160)	(0.142)	(0.496)	(0.497)	(0.496)
Number Pending Antitrust Cases Against Intel		0.015*			0.015*			0.015*			0.014*	
		(0.008)			(0.008)			(0.008)			(0.008)	
Number of Antitrust Cases Against Intel			0.014**			0.012*			0.014**			0.014**
			(0.007)			(0.007)			(0.007)			(0.007)

Notes: Robust Clustered Standard Errors in Parentheses. *** denotes significance at the 1% level, ** at the 5% level; and * at the 10% level. For space reasons we don't report parameter estimates for the initial conditions, standard deviations of the fixed effect, idiosyncratic error terms, quarterly trend, or constant. These are available upon request. Each regression contains 3108 observations.

3.8 Conclusion

Our findings reflect the importance of dynamics in the technology adoption process: downstream customers weigh the potential benefits from adopting AMD's technology against its costs, reflected in losing benefits from the Intel Inside program. Evaluation of both the benefits and the costs crucially depends on firm's expectations regarding several strategic variables: AMD's production capacity and technological progress contrasted with that of Intel, and the future viability of Intel's vertical restraints. The empirical evidence is therefore consistent with a nuanced form of foreclosure: one not targeted at eliminating a rival, but rather at keeping the rival below a certain threshold of production. In innovative industries that require substantial investments, this may successfully block the rival from developing the capacity and size that would allow it to threaten the incumbent's dominant position.

References

1. Arellano, M. and S. Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58: 277-97.

2. Arellano, M., and O. Bover (1995) "Another look at the instrumental variable estimation of error-components models", *Journal of econometrics*, 68(1), 29-51.

3. Asker, J. (2004) "Diagnosing foreclosure due to exclusive dealing", SSRN Working paper 609162.

4. Ater, I. (2015) "Vertical Foreclosure Using Exclusivity Clauses: Evidence from Shopping Malls", *Journal of Economics & Management Strategy*, 24(3), 620-642.

5. Berry, S., J. Levinsohn and A. Pakes (1995) "Automobile prices in market equilibrium," *Econometrica* 63(4), 841-890.

6. Besanko, D., and M. K. Perry (1993) "Equilibrium incentives for exclusive dealing in a differentiated products oligopoly", *The RAND Journal of Economics*, 646-667.

7. Blundell, R., and S. Bond (1998) "Initial conditions and moment restrictions in dynamic panel data models", *Journal of econometrics*, 87(1), 115-143.

8. Chamberlain, G. (1984) "Panel Data," in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Volume 2., 1247-1318.

9. Eizenberg A. (2011) "Upstream Innovation and Product Variety in the United States Home PC Market,"Working Paper.

10. Goettler, R. and B. R. Gordon (2011) "Competition and innovation in the microprocessor industry: Does AMD spur Intel to innovate more?" Working paper, Columbia University, New York.

11. Gordon, B. (2009) "A Dynamic Model of Consumer Replacement Cycles in the PC Processor Industry," Unpublished Working Paper, Columbia Business School.

12. Inderst, R., and Shaffer, G. (2007) "Retail mergers, buyer power and product variety", *The Economic Journal*, 117(516), 45-67.

13. Kaplow, P. A. L., P. Areeda, and A. Edlin (1997) "Antitrust analysis".

14. Lafontaine, F. and M. Slade (2008) "Exclusive Contracts and Vertical Restraints: Empirical Evidence and Public Policy" in Paolo Buccirossi, ed., *Handbook of Antitrust Econonmics*, Cambridge MIT Press, 2008, pp. 391-414.

15. Lee, H., Pechy, A., and M. Sovinsky (2013) "Is Intel's Marketing Campaign Predatory?," Working Paper, University of Zurich.

16. Lommerud, K. E., O. R. Straume and L. Sorgard (2005) "Downstream merger with upstream market power", *European Economic Review*, 49(3), 717-743.

17. Nurski, L., and F. Verboven (2016) "Exclusive dealing as a barrier to entry? Evidence from automobiles", *The Review of Economic Studies*, rdw002.

18. Sass, T. R. (2005) "The competitive effects of exclusive dealing: Evidence from the US beer industry", *International Journal of industrial Organization*, 23(3), 203-225.

19. Slade, M. E. (2005) "Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition," Working Paper.

20. Song, M. (2007) "Measuring Consumer Welfare in the CPU Market: An Application of the Pure Characteristics Demand Model," *RAND Journal of Economics*, 38, 429-446.

21. Sovinsky Goeree, M. (2008) "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 1017-1074.

22. Spector, D. (2011) "Exclusive contracts and demand foreclosure", *The RAND Journal of Economics*, 42(4), 619-638.

23. Sullivan, E. T. and H. Hovenkamp (1999) "Antitrust Law, Policy and Procedure", Lexis Law Publishing, Virginia.

24. Suzuki, A. (2009) "Market foreclosure and vertical merger: A case study of the vertical merger between Turner Broadcasting and Time Warner", *International Journal of Industrial Organization*, 27(4), 532-543.

25. Wooldridge, J. M. (2010) "Econometric analysis of cross section and panel data", MIT press.

26. Yehezkel, Y. (2008) "Retailers' choice of product variety and exclusive dealing under asymmetric information", *The RAND Journal of Economics*, 39(1), 115-143.

Appendix 3.A

Advertisement Variables We describe the creation of the advertisment variables out of the Kantar data. We identify 3 types of ad expenditures in the Kantar data: PC brand level advertising (Ad1), PC advertising categorized as business-to-business (Ad2) and PC firm level promotions (Ad3). We define two advertisment variables: brand specific advertising and firm level advertising. For brand specific advertising, we create the variable differently depending on whether the observation is in the home or non-home segment. Indeed, while Ad1 expenditures are likely to influence choices on both segments (households or firms), the Ad2 expenditures should only affect the non-home segment. The firm level advertising are identically defined on both segments and consist of Ad3. The definitions of the two variables are summarized in Table 3.A.1. For those observations of the advertisement data whose brand could not be matched with the Gartner data, the expenditures were accounted for as firm level expenditures. Finally, the above described ad variables were matched to the Gartner data at the brand level. As the Kantar data contains less details about PC brands than the Gartner dataset, the match occurred based on the Kantar brands.

Table 3.A.1 Segment-specific Definition of Advertisment Variables

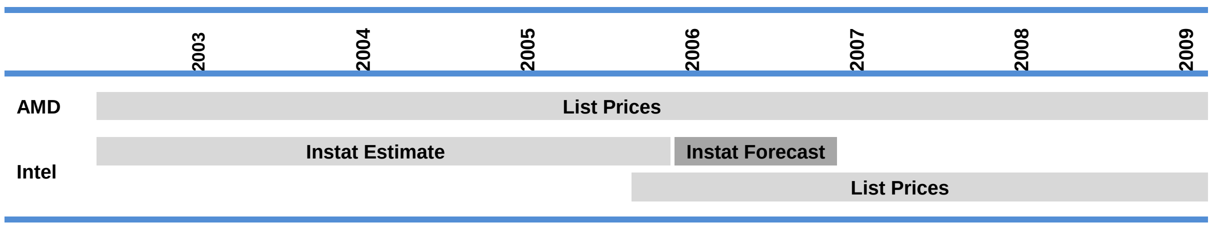
Variable / Segment	Home Segment	Non-home segment
Brand Advertising	Ad1	Ad1+Ad2
Firm Advertising	Ad3	Ad3

CPU Quality We measure CPU quality in terms of CPU benchmark per dollar. CPU benchmark information is gathered from Passmark publications.[18] This company collects measurements on CPU tests from users around the world, and creates a database of CPU performance at the CPU model level. Concerning CPU prices, we create our own CPU price dataset, as to our knowledge, a comprehensive CPU price database for the US in the time period of interest is not

[18] Source: https://www.cpubenchmark.net/

available. We use three different sources: Instat CPU core prices (estimates and forecasts), Intel list prices and AMD list prices. Table 3.A.2 offers an overview of the different data sources that we use and their respective coverage.

Table 3.A.2 Coverage of Various Price Data Sets



We first describe the creation of the CPU price dataset at the CPU family-quarter level. Intel prices for 2002Q3-2005Q4 are computed based on information from Instat's "Rosetta Stone" report on CPU core prices. We follow the methodology described in Lee, Pchy and Sovinsky (2013). A given CPU core is often marketed under different family names depending on which features are dis/enabled. For example, the CPU core "Northwood" is used in both "Pentium 4" and "Mobile Celeron" CPU families. Moreover, the CPU core used in a CPU family can change over time. Taking these into consideration, the CPU cores are matched to the CPU families of the PC data at the platform group (whether deskptop or mobile)/type (mainstream/value/ultraportable)/ family/speed/quarter level.[19] Table 3.A.3 provides the product cross-referencing.

[19] For the CPUs not matched at first attempt, the type is dropped. When unmatched, the data are matched based on family/marketing name of a CPU, CPU speed, year, and quarter, ignoring platform group. When the data are not matched, we try matching based on platform group, family/marketing name of a CPU, CPU speed, ignoring time. For observations still not matched, we take the averages of prices estimates of CPUs of the same marketing name, year and quarter.

Table 3.A.3 Product Cross-Reference from CPU Core to Family Name (i.e. Marketing Name) in 2002Q3-2005Q4

Platform		CPU Core	Family Name	Speed (Frequency: MHz)
Desktop	Mainstream	Willamette	Pentium 4	1300 - 2000
		Northwood		1600 - 3400
		Prescott		2260 - 3800
		Smithfield*	Pentium D	2667 - 3200
	Value	Tualatin	Pentium III	1000 - 1400
			Celeron	900 - 1400
		Willamette	Celeron	1500 - 2000
		Northwood		1600 - 2800
		Prescott	Celeron D	2133 - 3460
Mobile	Mainstream	Northwood	Mobile Pentium 4-M	1200 - 2600
		Prescott	Mobile Pentium 4	2300 - 3460
		Banias	Pentium M	1200 - 1800
		Dothan		1300 - 2267
	Value	Tualatin	Mobile Celeron	1000 - 1330
			Mobile Pentium III-M	866 - 1333
		Northwood	Mobile Celeron	1400 - 2500
		Banias	Celeron M	1200 - 1500
		Dothan		1200 - 1700
	Low-Power	Tualatin LV	Mobile Pentium III-M	733 - 1000
		Tualatin ULV		700 - 933
		Tualatin LV	Mobile Celeron	650 - 1000
		Tualatin ULV		650 - 800
		Banias LV	Pentium M	1100 - 1300
		Banias ULV		900 - 1100
		Dothan LV		1400 - 1600
		Dothan ULV		1000 - 1300
		Banias ULV	Celeron M	600 - 900
		Dothan ULV		900 - 1000

Notes: * Dual-core processor
Low-power mobile PCs are mini-notebook, tablet, and ultraportables.
(LV: low-voltage; ULV: ultra-low-voltage)

For the period 2005Q4-2009Q1, Intel prices were collected in the form of Intel's price catalogues (for 1000 units) from a large variety of websites. AMD prices for 1000 units were collected from the corporate website list price publications using waybackmachine.com, a website storing (many) historical saves of given websites. The list price information for both Intel and AMD are available at the CPU model-month level in variable frequency. This is due to irregular publication dates of the firms and, in a few cases , limited coverage of our sources. First, CPU prices at the month-model level are aggregated at the quarter-family level by taking the median over models. Second, prices are merged with the market share data at the family quarter level, to verify price data availability for each quarter of a family's market share sequence. For the periods of a sequence where price data is not available, we proceed as follows. When the price is missing in the middle of the sequence, it is approximated with kernel density interpolation at the family level. For prices missing in the first quarters of the sequence, the first observed price is used. These new introductions have usually very small market shares and high prices, which are preserved by this approximation. For prices missing in the last periods of the sequence, the last observed price is used. We obtain from the list prices, a dataset at the family quarter level with the following coverage: Intel 2005Q4 until 2009Q1, and AMD 2002Q3 until 2009Q1.

At this point, the Intel price data stem from two different sources: Instat for 2002Q3-2005Q4 and list prices for 2005Q4-2009Q1. To obtain a consistent measure of CPU prices, we define a correction coefficient. We take the median of the differences in percentage between CPU families' Instat and list prices for periods where both types of prices are available. As this is only fulfilled in period 2005Q4, we propose a second correction coefficient on periods 2005Q4-2006Q4 using the Instat CPU core price predictions for year 2006. The cross-referencing between Gartner and Instat is executed as previously described except for speed information, which is not available in the Gartner data for 2006. As for some CPU families more than one core is matched, the mean price over cores is retained. The cross-referencing is provided in Table 3.A.4. We obtain two different price measures for Intel CPUs based on these two correction coefficients (first over 2005Q4, second over 2005Q4-2006Q4). For robustness, we have run our regressions using each types of price measures.

Table 3.A.4 Product Cross-Reference from CPU Core to Family Name (i.e. Marketing Name) in 2006Q1-Q4

Platform		CPU Core	Family Name
Desktop	Mainstream	Conroe*	Celeron
		Conroe*	Core 2 Duo
		Prescott	Pentium 4
		Presler*	Pentium D
		Gallatin	Xeon
	Value	Cedar Mill	Celeron D
		Cedar Mill	Pentium 4
		Prescott	Celeron D
Mobile	Mainstream	Yonah*	Core Duo
		Dothan	Pentium M
	Value	Dothan	Celeron M
		Yonah	Celeron M
		Yonah	Core Solo
	Low-Power	Dothan LV	Pentium M
		Dothan ULV	Celeron M
		Dothan ULV	Pentium M
		Yonah LV	Xeon
		Yonah ULV	Celeron M
		Yonah ULV	Core Solo

Notes: * Dual-core processor
Low-power mobile PCs are mini-notebook, tablet, and ultraportables.
(LV: low-voltage; ULV: ultra-low-voltage)

We now discuss the treatment of the CPU benchmark information. The benchmark level of a given CPU family in a given period is built in two different ways for observations before and after 2005. Before 2005, CPUs of the Gartner sales data are matched to the Passmark data based on CPU family, CPU speed and platform information. Since after 2004 speed information is not available in the Gartner sales data, the availability of CPU models over time is inferred from our list price dataset described above. The matches between the CPUs of the benchmark and the list price data are achieved by taking the best of 3 different matching criteria (in order of preference):

family/model code/speed, family/model code, family/speed.[20] To obtain the level of observation, a CPU family quarter, we take the median of the benchmark level over CPU models in each quarter.[21] Finally, the variable of interest, the benchmark per dollar for AMD (resp. Intel) at the PC brand-segment level is computed in the following way. For each CPU family quarter, the ratio of benchmark per price is computed. Then the data is merged with the Gartner sales data, and we compute the average of this ratio weighted by market share of all PC models in a PC brand-segment. For a PC brand segment quarter where no AMD equipped PCs (resp. Intel) are offered, the benchmark per dollar variable is set to zero.

[20] Note that this last criteria is required in a minority of cases only. It can potentially aggregate very different benchmark levels (aggregating benchmarks of CPUs available in 2005 with some of 2008) to exclude these cases, we only use observations where the min and the max benchmarks are distant by less than 10%.

[21] For observations where benchmark information is missing, we use the same procedure as described above for prices (interpolation, first observed benchmark, last observed benchmark) since the benchmark data availability is corresponding to price data availability.

Chapter 4

Assessing an Efficiency Defense: The Case of Intel's Marketing Campaign

This chapter is co-authored with Hwa Ryung Lee and Michelle Sovinsky.

Acknowledgements: We thank Chloé Michel for excellent research assistance. We wish to thank Heski Bar Isaac, David Byrne, Alon Eizenberg, Yang Li, Carlos Noton, Minjae Song, Steve Stern, and Otto Toivanen for helpful comments and suggestions. Moreover, we thank seminar participants at the Conference on the Economics of Advertising and Marketing (Tel Aviv), ETH Zurich, Korea Development Institute, Korea University, MaCCI Summer Institute, Toulouse, the University of Washington, the University of Warwick, the Swiss IO Day and the University of Zurich.

4.1 Introduction

Generally, predatory pricing is a price reduction that is profitable only because of the additional market power gained from excluding or otherwise inhibiting the rival from competing. However, the predator may also induce rivals to exit the industry via non-price predation. Predatory investments could be made in excessive capacity, product differentiation, advertising, etc. For example, excessive investments that have the objective and likely effect of weakening or eliminating competitors can be predatory. Indeed in many predatory situations, pricing is only one aspect of anticompetitive behavior.

We present a framework to test if firms are using marketing/advertising campaigns in an anticompetitive fashion. Our "Test of Advertising Predation" (TAP) is based on the presumption that, if a firm's marketing campaign is not predatory, marketing expenses should be efficient (i.e., profit maximizing) and so should result in sufficient increased product demand to justify costs[1]. The TAP test examines if the return on advertising (i.e., how it impacts demand) is high enough to justify marketing expenditures (as these are directed at increasing demand). The test is based on a structural approach that allows us to disentangle the potential positive impact of a marketing program from potential anticompetitive predatory effects. Constructing the TAP test does not require any more "fancy" econometrics than that needed to estimate a model of demand. We use this test to examine if the Intel Inside marketing program, which provided marketing support for firms that sold Intel processors, was used in a predatory fashion (during 2002-2005). TAP results suggest short-term profit sacrifice by Intel over this period, which indicates that the Intel Inside campaign was used for predation.

While there is a vast theoretical literature on predation there are relatively few empirical studies, and these focus exclusively on pricing predation. Related papers in the price predation literature include: Weiman and Levin (1994) who examine predatory behaviour by Southern Bell Telephone Company between 1894 to 1912 when independent phone companies were trying

[1] Our test is related to advertising but as advertising is (a very important) marketing tool, we use the terms marketing and advertising interchangeably, while realizing that marketing can involve more than just advertising (e.g., corporate training).

to enter the market. Granitz and Klein (1996) provide evidence that Standard Oil engaged in predatory behavior by threatening to withhold inputs from railroads that were not in the railroad cartel. Genesove and Mullin (2006) estimate the price-cost margin in the sugar industry. They find that the price-cost margin was negative during price wars and predation was profitable in that it established a reputation as a tough competitor. Ellison and Ellison (2011) examine entry deterrence behavior in the pharmaceutical market prior to patent expiration by focusing on the asymmetry in detailing activities in markets of different size. Similarly, Chen and Tan (2007) focus on whether detailing in the pharmaceutical industry is consistent with predation incentives. Finally, Snider (2009) and Besanko, Doraszelski and Kryukov (2010) estimate dynamic models of predatory pricing.

Our work contributes to the recent stream of research that uses models to study strategic behaviour in the market for central processor units (CPU) and personal computers (PC). These include papers by Salgado (2008a), Song (2007), Gordon (2009), and Goettler and Gordon (2009) who study the upstream CPU market. This literature mostly abstracts from PC manufacturer and PC characteristics when estimating CPU demand, and assumes that final consumers buy CPUs directly. Instead we model consumer's choice of a PC and use it to infer CPU demand. An advantage of our approach is that we can more easily estimate the effect of advertising by a PC firm on demand. Given that the Intel Inside program is the marketing subsidy from Intel to PC firms, this will allow us to estimate the effect of the Intel Inside on demand more directly. As a result our work is related to the literature on estimating demand in the PC industry. Papers in this area include Eizenberg (2011), Sovinsky Goeree (2008), Prince (2008), and Gowrisankaran and Rysman (2007). Finally, we estimate the impact of advertising on PC demand, which is related to work by Sovinsky Goeree (2008) and Salgado (2008b).

This chapter is structured as follows. We describe the TAP test and explain how to implement it in section 4.2. We discuss why the Intel Inside campaign is a useful application of the TAP test and discuss the data we will use to conduct the test in section 4.3. In section 4.5, we develop and estimate a model of demand and present the parameter estimates. In section 4.6, we compute the marginal revenue of the market campaign, and use this to construct the tests presented in section 4.7. In section 4.9, we provide robustness tests and model specifications to address limitations of the TAP test. In the final sections we note the policy implications of this test and provide our concluding thoughts.

4.2 Test of Advertising Predation (TAP)

Predation is not a sensible business strategy if it cannot drive a rival out of a market, discipline a rival not to compete against a predator, or if the predator cannot maintain market power for a sufficient period of time after predation. Predation can be seen as an investment in long-run market power, and, as such we propose that marketing/advertising predation has two components: short-term profit sacrifice and long-term recoupment. We develop a "Test of Advertising Predation" (TAP) based on the presumption that, if the marketing/advertising campaign is not predatory, marketing expenses should be efficient (i.e., profit maximizing), and, so should result in sufficient increased product demand to justify costs.

The first step to construct TAP is to estimate product demand as a function of marketing variables. The second step is to use the demand estimates to compute firm marginal revenue derived from the marketing campaign. That is, one should compute the marginal revenue of marketing/advertising dollars spent on the marketing campaign (at the firm or product level). Notice that the marginal revenue of the marketing program depends on the parameters of consumer utility (including advertising/marketing variables), price and marginal manufacturing cost. Given data on marketing costs, the final step of TAP is to compare the marginal marketing revenue to the observed marketing costs. When observed marketing marginal costs are above the 95 percent confidence interval for the estimated marketing marginal revenue, we conclude that the marketing program is not consistent with profit maximization and, more specifically, that there was an excessive marketing subsidy. In this case the TAP test result is "positive".

Alternatively, the efficiency of the marketing program could be estimated by explicitly modelling how the marketing program works and estimating the parameters of the profit function. However, such an approach would have some non-trivial complications given that the way marketing programs function differs across firms and due to the fact that joint estimation of demand, supply, and optimal marketing choices would quickly complicate the econometrics. Furthermore, the resulting test would be program specific, and hence not applicable in many situations. A benefit of the TAP is that it allows us to circumvent modelling a firm's profit maximization problem (and hence their marketing program) while still allowing to compute an efficiency test.

4.2.1 Step 1: Estimate Demand

The first part of the test requires the researcher to specify a model of demand as a function of marketing variables. Notice that the demand specification can take any form as long as the marketing variables may have an impact on demand. Some common demand specifications include logit based perhaps with different levels of nests (e.g., Goldberg, (1995)) or random coefficients (e.g., Berry, Levinsohn and Pakes (1995) (hereafter BLP); Nevo (2001)). The benefits and drawbacks of these various specifications are well documented in the empirical IO literature. The demand specification chosen depends on a number of considerations including, the data that are available, the time available to conduct the TAP test, as well as the nature of the market being considered.

In addition to choosing a product demand system, it is important to consider how the marketing variables may impact demand. Some items to consider include the nature of the impact of marketing on demand. For example, is advertising assumed to influence the choice set of consumers (e.g., Sovinsky Goeree, (2008)), provide additional utility to the consumed good (e.g., Ackerberg (2003)).

For the sake of illustration, we assume that only aggregate data are available and specify product specific market shares (this would arise from a random coefficients BLP type demand, for example). Specifically, let $s_j(p, m \mid \Theta)$ represent the predicted market share for product (j), as a function of a vector of all product prices (p) and a vector of all product marketing variables

(m), where Θ represents the parameters to be estimated. If the market is one with multiproduct firms then firm f would sell the subset of products denoted by $\mathcal{J}_f$.

4.2.2 Step 2: Compute Marginal Marketing Revenue

The second step uses the demand side estimates to compute firm f's marginal revenue from the marketing campaign. In other words, this step computes the marginal revenue of marketing/advertising dollars spent on the campaign (this can be done at the firm or product level). The marginal revenue of the marketing program depends on the parameters of consumer utility (including marketing, price and marginal manufacturing cost, denoted mc_j).

Suppressing time notation, the total marketing/ad revenue (TMR) of firm f is given by

$$TMR_f = \sum_{j \in \mathcal{J}_f} (p_j - mc_j)\mathcal{M}s_j(p, m),$$

where $\mathcal{J}_f$ is the set of firm f's products; $\mathcal{M}$ is the market size; $s_j(p, m)$ is market share, which depends on product prices in the market (p) and marketing (m). Firm f's marginal revenue from the marketing campaign (MMR) is given by

$$MMR_f = \mathcal{M} \sum_{r,j \in \mathcal{J}_f} (p_j - mc_j)\frac{\partial s_j(p, m)}{\partial m_r},$$

where a multiproduct firm will consider the impact on its complete product line arising from a change in marketing for its product r. One important caveat: to compute the marginal marketing revenue it is necessary to have data on (or an estimate of) the marginal cost of production.

4.2.3 Step 3: Compare to Marginal Cost of Marketing Campaign

The final step of TAP is to compare the marginal marketing revenue to the observed marketing costs (MC). However, in most situations the researcher will not observe marginal marketing costs.

To overcome this data restriction a solution implemented in the predatory pricing literature is to use average variable costs to proxy for marginal costs.[2] In this setting, the researcher does not require a proxy for marginal costs of production, but rather a proxy for marginal marketing costs. However, the same solution can be applied to use average variable marketing costs as a proxy for marginal marketing costs.

There may be a concern whether the average cost of the marketing program is a proper measure of the marginal cost of the marketing program. This issue arises in the predatory pricing literature as there is a practical difficulty in determining the nature of production cost: whether it is variable or fixed, or whether it is a sunk cost. However, this is less of a concern as it relates to marketing as there are few fixed or sunk components to advertising expenditures. Aside from practical difficulties, average variable cost may not be a good proxy for marginal marketing cost in the presence of returns to scale of the marketing program. The TAP results will reflect the assumption that firms' responsiveness to the marketing program is constant.

4.2.4 Step 4: Consider TAP Limitations

There are few issues to consider when using the TAP efficiency test. First, TAP is fundamentally a test of profit sacrifice. Hence, a case would need to be made that long-term profit recoupment is possible. Second, short-term profit sacrifice can be rationalized by potential dynamic efficiency reasons such as learning-by-doing, promotional purposes, or network externalities. As the TAP test examines if the return on advertising (i.e., how it impacts demand) is high enough to justify marketing expenditures (as these are directed at increasing demand) the model includes only the current, short-term effect of advertising. Hence, the potential long-run benefit of the marketing program is not taken into account. Note this is less of a concern when considering firms that

[2] There are other proposals for how to calculate marginal costs associated with predatory pricing behavior. For example, Bolton, Brodley and Riordan (2000, 2001) suggest that the relevant cost is not average variable cost but the long run average incremental cost. This is measured by the per unit cost of producing the predatory increment of output where all costs that were incurred (regardless of when they were incurred) are considered. Specifically, it is calculated as the firm's total production cost less what the firm's total cost would have been had it not produced the predatory units divided by the quantity of the product produced. There is no analogy for the advertising predation measure that we could construct without measuring the output produced under the predatory behavior.

have been operating for a while: just as the efficiency reasons for pricing below marginal cost are not usually applicable to an already dominant, incumbent firm with a large customer base, in this setting too an unprofitable advertising/marketing subsidy is not easily justified by efficiency reasons.

A third issue relates to the brand-loyalty-building effect of marketing programs. Advertising is generally believed to build goodwill and this may be a reason to invest in marketing at the expense of short-term profits. Notice that this critique is less important if there is some time element during which the marketing program behavior is being investigated. The reason being that the incentive to build goodwill is constant across all periods but the predatory motive would only be present during certain periods. One way to examine the dynamic efficiency argument is to include lagged marketing variables in the demand specification. Another way to address this concern is to determine how many future periods of MR would need to be considered by firm to rationalize current period MCs.

4.3 Application of TAP to "Intel Inside" Campaign

4.3.1 Background on "Intel Inside" Campaign

Intel has been investigated for predatory (pricing), exclusionary behavior, and the abuse of a dominant position in the market for CPU. According to U.S. lawsuits, Intel used marketing loyalty rebates, payments, and threats to persuade computer manufacturers, including Dell and Hewlett-Packard (HP), to limit their use of AMD (Intel's main rival) processors. In their investigations, U.S. antitrust authorities focused on whether the loyalty rebates used by Intel were a predatory device in violation of the Sherman Act. The European Commission (EC) brought similar charges and imposed a 1.06 billion Euro fine on Intel for abuse of a dominant position. South Korean and Japanese antitrust authorities also imposed fines on Intel for breach of antitrust regulations.

In the case of Intel, an important component to the case involved their marketing campaign, "Intel Inside," which provided marketing support for firms that sold Intel CPU chips. Specifically, it is a cooperative advertising program in which Intel contributes a percentage of the purchase price of processors to a pool for PC firms to use to market Intel-based computers. According to the rules of the program PC firms can receive a rebate of their marketing expenditures if they include the Intel logo in their advertising. By the end of the 1990s, Intel had spent more than $7 billion on the marketing campaign (Moon and Darwall (2005)).

Intel was accused of using the marketing program to attempt to prevent computer makers from offering machines with non-Intel computer chips. It became clear through correspondence that Intel was trying to circumvent antitrust laws by using non-price predatory avenues. For example, a 2002 Dell document states that the "original basis for the [Intel marketing] fund is ... Dell's loyalty to Intel." The document explains that this means "no AMD processors."[3] The beginning of the alleged anticompetitive use of the Intel Inside program coincides with the introduction by their main rival AMD's Athlon chip (in 1999). Antitrust documentation shows that Intel issued "conditional rebates" from December 2002 to December 2005, whereby they would give rebates to some PC firms (Dell in particular) under the condition that the PC firm buy exclusively from Intel.[4] Otherwise, Intel would retract the marketing rebate and instead use the market development money to fund competitors. An internal Dell presentation (in 2003) noted that if Dell switched to AMD, Intel's retaliation "could be severe and prolonged with impact to all LOBs [Lines of Business].". Intel allegedly treated HP, Lenovo, and Acer similarly. For example, Intel rebates were conditional on HP buying 95% of its microprocessors for business desktops from Intel. In 2002, an HP executive wrote "PLEASE DO NOT . . . communicate to the regions, your team members or AMD that we are constrained to 5% AMD by pursuing the Intel agreement.".

We focus on Intel's marketing subsidy to Dell during the $2002-2005$ period to take advantage of antitrust documentation on marketing rebate payments made to Dell. Although the data are

[3] US District Court for the District of Delaware Complaint (2009)

[4] U.S. District of Court for District of Columbia; SEC (Securities and Exchange Commission) vs. Dell, pp. 10-11 and U.S. District of Court for District of Delaware; State of New York, by Attorney General Andrew M. Cuomo vs. Intel Corporation, p.6.

not as extensive for other PC firms, we evaluate the TAP test for firms involved the Intel antitrust case (HP and Toshiba) and a firm that was not involved (Gateway).

4.4 Data

We use three main data sources for our analysis: PC and CPU sales are from Gartner Group, advertising data are from Kantar Media Group, and CPU price and cost data are from In-Stat. All data are available from the first quarter of 2002 through the last quarter of 2005. We discuss each in turn.

Quarterly PC and CPU sales are at the product level, where a product is defined as PC vendor (i.e., Acer), PC vendor brand (i.e., Aspire), platform type (i.e., Notebook), CPU vendor (i.e., Intel), CPU family (i.e., Pentium 4), CPU speed (i.e., 1600/1799 MHz) combination. We focus on the market for US home consumers for two primary reasons. First, businesses make multiple purchases at a time, which would greatly complicate the empirical model, and, second, we don't have access to advertising data for each sector separately.

Advertising data consist of PC advertising expenditures.[5] The advertising data are quite detailed, sometimes even at the level of a specific product/model (e.g., Acer Aspire AS5735 Notebook Computer). However, it is difficult to match with the data from Gartner Group because the definition of products/models varies between the two datasets. Kantar Media Group uses a model name, such as Aspire AS5735; whereas, Gartner Group defines PC models as a combination of PC vendor, PC brand, platform type, CPU vendor, CPU family, and CPU speed. We match the two data sets based on brand and platform type. Table 4.1 shows the market share and total PC advertising expenditures by PC firm in the entire sample.

In-Stat provides data on CPU prices and manufacturing costs for selected processors and time periods. We need to match our PC data (where a CPU is in a PC) from Gartner group with CPU

[5] PC firms advertise printers and other computer accessories. We do not include these advertising expenditures.

prices and manufacturing costs from In-Stat. CPU prices are available by processor core

Table 4.1: Market Shares and Advertising Expenditures by PC Firms

PC firm	Num.obs	Market share (% shipment)	Quarterly Average Total PC-related advertising (M$)
Acer	428	0.31%	0.87
Apple	223	4.80%	-
Averatec	37	0.42%	0.00
Dell	1020	32.44%	1.58
emachines	59	3.86%	0.004
Fujitsu	193	0.30%	3.06
Gateway	487	11.62%	9.98
HP	1438	29.17%	65.92
IBM/Lenovo	535	0.23%	23.17
Sony	360	2.93%	5.08
Systemax	507	0.36%	0.19
Toshiba	294	3.54%	4.19
Other	1867	10.03%	-
Total	7,448	100%	

Notes: Market share is total firm PC Shipments / total industry PC shipments. *Our measure of advertising includes all PC-related advertising expenditures even w hen a target product w as not sold in a quarter.

on a quarterly basis[6]. In our PC data, we know the CPU family (that is, the marketing name, e.g., Pentium 4) and speed (frequency) of the CPU. The same processor core is often used to make processors that are marketed under different family names with different sets of features enabled, and the processor core used in a processor changes over time as technology advances. For instance, processor core "Williamette" was used for processor families marketed as Pentium 4 and as Celeron for desktop computers, while in later periods the same CPU families switched to the next-generation processor core "Northwood." We match the data based on platform group (whether desktop or mobile), type (whether mainstream, value, or ultraportable), family/marketing name of a CPU, CPU speed, year, and quarter.[7] We provide the product cross-reference in Table 4.A.1[8] in the appendix.

[6] CPU prices are available at several different levels of detail. The most detailed information is list prices of specific processors (e.g. Pentium M 1.40GHz). These prices are available for selected processors from July 2004 to July 2005, mostly on a monthly basis. Although it would be ideal to have list prices for all processors for all time periods, these detailed data cover only a subset of our sample.

[7] This process (and a slight generalization of it described below) generates a high match. For example, among Dell PCs, we have 78% match. For the CPUs not matched at first attempt, we drop type, then we have 83% match. When unmatched, the data are matched based on family/marketing name of a CPU, CPU speed, year, and quarter, ignoring platform group. Then we obtain a 96% match. When the data are not matched, we try matching based on platform group, family/marketing name of a CPU, CPU speed, ignoring time, and then we have 99% match. For observations still not matched, we take the averages of prices and cost estimates of CPUs of the same marketing name, year and quarter.

[8] The cross-reference table is constructed based on In-Stat's document and an website specialized in CPU information, www.cpu-world.com.

CPU manufacturing cost estimate data are more limited in that the cost estimates are available by CPU processor core for a broader definition as of 2005. For processor core "Williamette", we have cost estimates for different years, but not throughout the data period. Intel has constructed fabs and changed the use of existing fabs, which affected cost levels over time. Also, learning-by-doing drives the cost level downward over time and so cost depends on how mature the manufacturing process is. We use two approaches to construct our CPU marginal cost variable. First, we use In-Stat cost estimates matched with PC data using the cross-reference Table A1.[9] Table 4.2 presents the summary statistics for the price and cost estimates of Intel CPUs in the sample. Recall that the cost estimates are not time-varying for almost all processor cores.

Table 4.2: CPU price and CPU marginal cost of Intel-based Dell PCs

	Num.obs	Mean	Std.Dev.	Min	Max
CPU price	1020	146.98	57.14	49	317
CPU marginal cost	1020	36.64	5.77	26	57

Table 4.3 shows the percentage of PCs sold by PC firm and CPU vendor as well as the overall market share that the firm contributes toward the CPU manufacturers product over the sample period. As the Table illustrates Dell, IBM/Lenovo, and Toshiba used Intel based CPUs exclusively. Dell's purchases of Intel based CPUs contributed the most (32%) of any PC firm to Intel's market share, with HP following a close second (23%). These contributions to Intel's market share are significantly higher than the next closest PC firm, which is Gateway at 9%.

The summary statistics in Table 4.4 indicate, 88% of the PCs have an Intel CPU, where the average price of a PC is $1,250. Over half of the PCs are mobile as opposed to desk-based. Approximately 19% of the PCs were shipped in the first quarter of the data period, i.e., 2002:Q1 (these are denoted *Older PC*) and about 36% of CPUs where used in these PCs. [10]

[9] As for reliability of the cost estimates, In-Stat document states "Equations to calculate the number of die sites per wafer, yield, and cost per good die are well known throughout the industry. Important physical parameters, such as package type and die size, are generally published by the vendor and are verifiable through destructive analysis. The key areas of uncertainty are in estimating wafer cost, defect density, testing cost, and package

Table 4.3: Percent and Market Share of PCs by CPU Type

	Intel		AMD		Total
	% PCs	Market share	% PCs	Market share	Num.obs
Acer	89.72%	0.26%	10.28%	0.05%	428
Averatec	40.54%	0.13%	59.46%	0.30%	37
Dell	100.00%	32.44%	0.00%	0.00%	1020
emachines	57.63%	1.75%	42.37%	2.11%	59
Fujitsu	88.82%	0.28%	11.18%	0.01%	170
Gateway	93.84%	8.60%	6.16%	3.23%	487
HP	78.13%	23.03%	21.87%	6.14%	1436
IBM/Lenovo	100.00%	0.23%	0.00%	0.00%	535
Sony	90.40%	2.77%	9.60%	0.14%	354
Systemax	74.16%	0.29%	25.84%	0.07%	507
Toshiba	100.00%	3.54%	0.00%	0.00%	294
Total	88.38%		11.62%		5327

If the PC is not an *Older PC* then *PC age* indicates a mean of 1.6 quarters since the first shipment of the PC and 2.5 quarters since the first sales record of a CPU. The age variables (*Older PC*, *PC age*, *Older CPU*, and *CPU age*) are intended to capture the quality (how obsolete the production technology is), popularity and consumer awareness of a product (how long it has been on the market). CPU benchmark is a (continuous) quality benchmark that compares the relative speeds of different CPUs (collected by PassMark[11]). CPU manufacturers spent an average 23 million dollars on general firm promotions and chip advertising while PC firms spend an average 4.4 million dollars for PC brand advertising. For the summary statistics we present the market share weighted measure of PC brand advertising.[12]

cost."

10 If there is a shipment record of a CPU in the first quarter of the data period, we can assume that the CPU has been introduced in that quarter or earlier.

11 CPU Benchmark results were gathered from users' submissions to the PassMark web site (http://www.cpubenchmark.net/cpu_list.php) as well as from internal testing. PerformanceTest conducts eight different tests and then averages the results to determine the CPU Mark for a system. To ensure that the full CPU power of a PC system is realized, PerformanceTest runs each CPU test on all available CPUs. Specifically, PerformanceTest runs one simultaneous CPU test for every logical CPU (Hyper-threaded); physical CPU core (dual core) or physical CPU package (multiple CPU chips). So hypothetically if you have a PC that has two CPUs, each with dual cores that use hyper-threading then PerformanceTest will run eight simultaneous tests. Since PassMark point is not available for some models, we use a linear interpolation for those missing data, based on CPU model and CPU speed.

12 We construct a brand-platform advertising variable with a market share weight. It is the sum of brand-platform advertising, brand advertising weighted by market share of a platform within the brand, brand combination advertising weighted by market share of brand-platform within the brand combination, platform advertising weighted by market share of brand within the platform, and firm advertising weighted by market share of

Table 4.4: Descriptive Statistics

Variable	Mean	Min	Max
Price (1000$2000)	1.25	0.38	3.52
Intel CPU	0.88	0	1
Mobile PC	0.57	0	1
CPU benchmark	0.34	0.13	0.9
CPU speed (100mhz)	2.08	0.65	3.8
Older PC	0.19	0	1
PC age	1.6	0	10
Older CPU	0.36	0	1
CPU age	2.52	0	10
Chip ads (10 mil.$2000)	2.39	0.02	4.66
PC brand ads (10 mil.$2000)	0.44	0	6.49
Num. obs.	5327		

Notes: Price (unit: 1000$) and advertising (unit: 10 million $) variables are adjusted to 2000 dollars using Consumer Price Index (CPI) data from the U.S. Department of Labor Bureau of Labor Statistics.

Table 4.5 presents descriptive statistics for the two main CPU manufacturers: Intel and AMD. As the table indicates, over the data period, average AMD chips used in PCs performed significantly better than average Intel chips (see *CPU benchmark*). On average, we can see that the PC models with Intel chips are priced higher and advertised significantly more at the CPU level. In many cases, the same PC brand has models with AMD chips and models with Intel chips. Thus, PC brand ads do not accurately capture the difference between PCs with Intel chips versus those with AMD chips. Although PC brand ads are larger for the brands of PCs with AMD chip on average, it is because some heavily advertised brands have models with AMD and Intel chips.

brand-platform within the firm. Market shares are computed based on shipments. For model j of brand b, platform p, and firm f, we have:

$$\begin{aligned} ad_j^{\text{brand}} &= a_{b,p}^{\text{brand,platform}} + \frac{\sum_{i\in\mathcal{J}_{fpb}} s_i}{\sum_{i\in\mathcal{J}_{fp}} s_i} a_{b,f}^{\text{brand}} + \frac{\sum_{i\in\mathcal{J}_{fpb}} s_i}{\sum_{b'\in\mathcal{J}_{fpc}} \sum_{i\in\mathcal{J}_{fpb'}} s_i} a_{c,p}^{\text{brand comb}} \\ &+ \frac{\sum_{i\in\mathcal{J}_{fpb}} s_i}{\sum_{i\in\mathcal{J}_{fp}} s_i} a_{f,p}^{\text{platform}} + \frac{\sum_{i\in\mathcal{J}_{fb}} s_i}{\sum_{i\in\mathcal{J}_{f}} s_i} a_{f}^{\text{firm}} \end{aligned}$$

where $\mathcal{J}_*$ denotes the set of models in category $*$, $a_{b,p}^{\text{brand,platform}}$ is brand-platform advertising, $a_{b,f}^{\text{brand}}$ is brand advertising; $a_{c,p}^{\text{brand comb}}$ is brand combination advertisin, $a_{f,p}^{\text{platform}}$ is firm-platform advertising, and a_{f}^{firm} is firm advertising. This is a similar methodology that is used in the literature for quantity weighted average prices (see for example, Song (2007)).

Table 4.5: Descriptive Statistics by CPU Manufacturer

	Mean Intel	Mean AMD	Difference	t-value
Price (1000$2000)	1.28	0.98	0.30	17.10
Mobile	0.58	0.52	0.06	2.63
CPU benchmark	0.33	0.46	-0.14	-26.01
CPU speed (1000mhz)	2.09	1.93	0.16	5.20
Older PC	0.19	0.21	-0.02	-1.34
PC age	1.65	1.25	0.40	4.74
Older CPU	0.37	0.33	0.04	1.87
CPU age	2.60	1.93	0.67	5.64
Chip ads (10 mil.$2000)	2.65	0.36	2.29	45.79
PC brand ads (10 mil.$2000)	0.41	0.69	-0.28	-8.35
Num. obs.	4708	619		

Finally, we use surveys on PC purchases from Forrester Research from 2002 through 2005. These data have information about individual consumers' PC and CPU choices, although they are not detailed at the product level. For example, we observe whether a survey respondent bought a PC in the last year, some characteristics of the PC such as PC firm and CPU manufacturer (Intel, AMD, Apple[13], Other, or Don't know) if purchased.

4.5 Demand for CPUs

The demand for CPUs can be inferred from the demand for a PC model as a PC comes equipped with a single CPU. When consumers purchase computers, they choose a combination of PC firm and CPU type.[14] There are T markets, indexed by $t = 1, 2, ..., T$. A home market consumer chooses from J products, indexed $j = 1, ..., J$, where a product is a PC vendor (i.e., Acer), PC vendor brand (i.e., Aspire), platform type (i.e., Notebook), CPU vendor (i.e., Intel), CPU family (i.e., Pentium 4), CPU speed (i.e., 1600/1799 MHz) quarter combination. Product j characteristics

[13] During the data time period, Apple PCs used only IBM chips.

[14] Many websites which provide CPU performance comparisons, such as CPUScoreCard.com, categorize CPUs depending on computer type (i.e., desktops or laptops), and they compare CPUs within the same computer type. Considering that CPU chips intended to be used in desktops and laptops are different due to different requirements, we can think that the choice of CPU type includes the choice of computer type. Song (2007) and Salgado (2008b) modeled consumers' choice of CPU type although consumers more often buy computers, rather than CPU chips without computers.

are price of the PC (p) and non-price observed attributes of the PC (x), which include the platform and PC vendor dummy variables, dummy variables for whether the processor is manufactured by Intel, processor speed, the age of the CPU and the CPU benchmark score.

Advertising is an additional characteristic that may impact consumer demand.[15] We observe various levels of aggregation of advertising expenditures by PC firms. These include firm-specific advertising (i.e., advertising for Dell), firm-brand specific advertising (i.e., Dell Presario), firm-platform advertising (i.e., Dell Notebooks), and firm-brand-platform advertising (i.e., Dell Presario Notebooks). CPU vendor advertising is at the firm level (i.e., Intel) and at the CPU level. Given the difference in advertising campaigns across firms and CPU suppliers, we allow advertising by PC firms to have different effect on consumer utility than advertising done by CPU vendors.[16] The a_{jt}^{pc} is a vector of PC advertising variables aggregated at different levels and a_{jt}^{cpu} is a vector of CPU advertising variables aggregated at different levels. The ϵ_{jt} is a mean zero stochastic term i.i.d. type I extreme value distributed across products and markets. Consumers have an "outside" option, which includes purchase of a computer with non-Intel or non-AMD processor (such as IBM chips exclusively used by Apple computers during the sample period)[17], a PC manufactured by a small firms[18], and self-assembled PCs. The indirect utility of this outside option is normalized to zero.

The indirect utility a consumer obtains from j at time t is

$$u_{jt} = p_{jt}\alpha + x_{jt}\beta + a_{jt}^{pc}\gamma + a_{jt}^{cpu}\lambda + \epsilon_{jt}.$$

The corresponding predicted market share can be written as:

$$s_{jt} = \frac{exp(u_{jt})}{1 + \sum_r exp(u_{rt})}. \tag{4.1}$$

15 It is reasonable to conjecture that all consumers know of the existence of Intel processors, therefore we do not model Intel advertising as impacting the consumer's choice set, but rather as impacting utility directly. We assume that consumers know all firms and processor types when making a purchase decision.

16 We also consider a specification in which we allow for nonlinear effects of advertising.

17 Apple Computer is included in the outside option for CPU's because they used IBM processors during the sample period (2002:Q2 - 2005:Q4).

18 These include Everex, Medion, Micro Electronics, Motion Computing, MPC, NEC, Sharp, and Velocity Micro.

And this implies the market share for firm f of processor type c is given by

$$s_{fct} = \sum_{j \in \mathcal{J}_c \cap \mathcal{J}_f} s_{jt}.$$

4.5.1 Preliminary Demand Estimates

First, we present results from a series of probit regressions of the probability of purchasing a PC in 2002 for use at home as given in the Forrester data. The estimates are presented in Table 4.6 and illustrate the importance of observed product heterogeneity in PC purchases. The results indicate that the form factor of the PC is important as well as firm fixed effects. Finally, the opportunity to add-on certain items (such as a flat-screen monitor) increases the probability of purchase. As columns (4)-(6) illustrate, consumers value PCs with a Intel or AMD processor even after controlling for other product characteristics. Furthermore, the results suggest the valuation of processor type is not constant over time (Column 6).

Table 4.7 presents probit estimates of purchase probabilities that illustrate the importance of product and individual observed heterogeneity for PC purchases. The independent variable is again whether the individual bought a PC in the past year. We start by including demographic variables (Column 1). We can see all the estimates on the coefficients of demographic variables are statistically significant. Once we include a laptop dummy and CPU type (Intel dummy and AMD dummy), some demographic variables become less significant (Column 2). Not surprisingly, the results indicate individuals are more likely to purchase a PC with an Intel or AMD processor than non-branded processors. Whether the computer is a laptop or not is also a significant factor in the purchase decision. After controlling for PC add-ons (Column 3) and Operating System dummy variables (Column 4), preference for a laptop and a PC with Intel or AMD processor have less impact on PC purchase probability. Overall, the results suggest that PC characteristics affect the purchase decision after controlling for demographics and that certain demographics are more important than others in the purchase decision.

Table 4.8 presents multinomial logit estimates of the choice of CPU manufacturer. The results show that there is significant consumer observed heterogeneity with respect to CPU choices. Given that a consumer purchased a PC in the past year, brand CPUs are preferred by a male, single, or higher-income household to non-brand CPUs. Among brand CPUs, the choice probability of a particular CPU varies with demographics. The presence of teenagers in the household makes purchase of an Intel or Apple less likely, but doesn't significantly impact the purchase of an AMD processor. Being young and being white are significantly and positively associated with the choice of Intel or AMD processors.

4.5.2 Demand Estimation Results

In Table 4.9 we present results from a Multinomial Logit PC-CPU demand model with aggregate market data, where only product characteristics are included as explanatory variables based on equation (4.1).[19] Advertising variables are aggregated at the brand-platform level. How we allocate advertising expenditure to PC brand-platform level is detailed in data section 4.4. We also include the squared value of the sum of general promotions and brand combination advertising that can be attributed to a given brand to allow for decreasing or increasing return to advertising. In every model specification, PC-firm-specific dummy variables and time-specific dummy variables are included.

The first specification column (1) illustrates the importance of observed product characteristics for PC/CPU choice. The results imply that processor speed and performance (benchmark) have a significant and positive effect on utility. This is reasonable as it suggests consumers prefer faster and better-performing CPU chips. However, the age of the CPU does not impact consumers valuations of the PC, while the age of the PC does. A likely explanation is that age captures the popularity or consumer awareness of a PC model or a CPU. If a product has been on the market for a while, it may imply either that consumers like the product, due to attractive qualities

[19] Estimates from a nested logit model (with laptop and desk-based PC nests) indicate that purchases do not take place in this nested structure in the sense that the estimated coefficient on the substitutability between the nests is often larger than one.

unobserved by the researcher, or that the product is well-known to potential buyers. Surprisingly, consumers place a lower valuation on the PC if it is a laptop, but this finding is not robust across specifications. Finally, conditional on CPU speed and the performance benchmark, consumers place a lower value on CPUs manufactured by Intel.

As we discussed earlier, prices may be correlated with the structural error term and hence endogenous. In column (2) we include instruments for price. The impact of price on demand becomes much more negative, which is consistent with price being a proxy for higher quality. The rest of the estimates do not change with the exception of the valuation of Intel and the valuation of laptop, which are no longer significant.

In specification (3), PC characteristics are taken out and advertising variables are added. As compared with the price coefficient estimate from specification (1), the price coefficient is estimated to be smaller (i.e. more negative). This suggests that advertising variables may be correlated with unobserved product attributes and we need to correct for the possible correlation between advertising and unobserved high quality. We also find that consumers marginal valuation of PC advertising is positive, while the valuation of firm level CPU advertising is negative and statistically insignificant.

In specification (4) we allow the effect of advertising to be different across PC firms. The results suggest some firms are more effective at advertising their products than others, but otherwise the coefficient estimates do not change much. Specifications (5) and (6) include all explanatory variables (PC characteristics, CPU characteristics, and advertising variables), but specification (5) instruments only for price while specification (6) instruments for both price and advertising. We have more elastic demand when using specification (6). Again, this suggests that advertising is correlated with unobserved product attributes.

Table 4.6: Probit Regressions of PC Purchase (in 2002)

Variables		(1)	(2)	(3)	(4)	(5)	(6)
		Dependent Variable: Whether Bought a PC in the past year					
Laptop		0.535***	0.473***	0.395***	0.527***	0.473***	0.474***
		(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)
Processor Manufacturer	Intel				0.139***	0.021***	0.071***
					(0.008)	(0.008)	(0.015)
	Intel and 2003						-0.056**
							(0.022)
	Intel and 2004						-0.085***
							(0.022)
	Intel and 2005						-0.061***
							(0.020)
	AMD				0.511***	0.344***	0.484***
					(0.016)	(0.016)	(0.028)
	AMD and 2003						-0.232***
							(0.044)
	AMD and 2004						-0.139***
							(0.045)
	AMD and 2005						-0.230***
							(0.041)
PC Manufacturer	Acer	-0.416***	-0.367***	-0.202***	-0.396***	-0.346***	-0.346***
		(0.042)	(0.043)	(0.045)	(0.042)	(0.043)	(0.043)
	Apple	-0.035*	-0.107***	-0.363***	0.093***	-0.055***	-0.050**
		(0.019)	(0.019)	(0.022)	(0.019)	(0.020)	(0.020)
	AST	-1.052***	-1.018***	-0.868***	-1.024***	-0.988***	-0.988***
		(0.166)	(0.167)	(0.174)	(0.167)	(0.168)	(0.168)
	Compaq	-0.056***	-0.089***	-0.125***	-0.045***	-0.078***	-0.078***
		(0.012)	(0.012)	(0.013)	(0.012)	(0.013)	(0.013)
	Dell	0.457***	0.350***	0.131***	0.486***	0.387***	0.386***
		(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
	Emachines	0.348***	0.311***	0.128***	0.371***	0.332***	0.334***
		(0.017)	(0.018)	(0.019)	(0.017)	(0.018)	(0.018)
	Fujitsu	-0.526**	-0.623***	-0.625**	-0.491**	-0.589**	-0.593**
		(0.231)	(0.238)	(0.255)	(0.231)	(0.238)	(0.238)
	Gateway	-0.163***	-0.253***	-0.317***	-0.135***	-0.222***	-0.223***
		(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
	HP	0.089***	0.021*	-0.093***	0.112***	0.046***	0.046***
		(0.011)	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)
	IBM	-0.279***	-0.259***	-0.169***	-0.247***	-0.232***	-0.232***
		(0.019)	(0.020)	(0.021)	(0.020)	(0.020)	(0.020)
	Sony	0.358***	0.096***	-0.114***	0.366***	0.120***	0.119***
		(0.025)	(0.026)	(0.027)	(0.025)	(0.026)	(0.026)
	Toshiba	0.195***	0.149***	0.055*	0.215***	0.173***	0.172***
		(0.027)	(0.028)	(0.030)	(0.027)	(0.028)	(0.028)
PC Add-ons	Broadband Adapter		0.229***	0.157***		0.223***	0.220***
			(0.012)	(0.012)		(0.012)	(0.012)
	DVD		0.447***	0.266***		0.435***	0.436***
			(0.008)	(0.008)		(0.008)	(0.008)
	CD Rom		0.044***	0.015		0.029***	0.029***
			(0.010)	(0.011)		(0.010)	(0.010)
	Flat Panel Monitor		0.581***	0.421***		0.580***	0.582***
			(0.010)	(0.011)		(0.010)	(0.010)
	Webcam		-0.128***	-0.184***		-0.140***	-0.136***
			(0.014)	(0.015)		(0.014)	(0.014)
Operating System	Windows 95 or 98			-0.918***			
				(0.014)			
	Windows ME			-0.195***			
				(0.016)			
	Windows 2000			-0.236***			
				(0.016)			
	Windows NT			-0.064*			
				(0.035)			
	Windows XP			0.442***			
				(0.014)			
Observations		167221	167221	167221	166246	166246	166246

Notes:Standard errors are in parenthesis. *** indicates significant at 1%; ** at 5% and * at 10%. All regressions include time dummies.

Table 4.7: PC purchase Probability on PC Characteristics and Demographics

Dependent Variable: Whether Bought a PC in the past year				
Variables	(1)	(2)	(3)	(4)
Laptop		0.518***	0.477***	0.409***
		(0.011)	(0.011)	(0.011)
Intel Processor		0.118***	0.024***	-0.018**
		(0.008)	(0.008)	(0.009)
AMD Processor		0.491***	0.351***	0.225***
		(0.016)	(0.017)	(0.018)
Demographics				
Age	-0.005***	-0.004***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.046***	0.014*	0.006	0.016**
	(0.007)	(0.007)	(0.007)	(0.008)
White	0.031**	0.028**	0.031**	0.003
	(0.012)	(0.013)	(0.013)	(0.014)
Married	0.037***	0.052***	0.039***	0.025**
	(0.009)	(0.009)	(0.009)	(0.010)
Presence of Teenagers	0.145***	0.156***	0.134***	0.113***
	(0.009)	(0.009)	(0.010)	(0.010)
Income > $100,000	0.176***	0.100***	0.047***	0.024**
	(0.009)	(0.010)	(0.010)	(0.010)
Income < $25,000	-0.083***	-0.003	0.046***	0.073***
	(0.008)	(0.008)	(0.008)	(0.009)
Fixed Effects				
Firm		Included	Included	Included
Year	Included	Included	Included	Included
PC-Add ons			Included	Included
Operating System				Included
Observations	163,187	160,982	160,982	160,982

Notes: Standard errors are in parenthesis. *** indicates significant at 1%; ** at 5% and * at 10%.

Table 4.8: CPU Purchase Estimates

Dependent Variable: CPU in a PC purchased in the past year (Base: Non-Intel/AMD/Mac or unknown)			
Variables	Intel	AMD	Apple CPU
Age	-0.018***	-0.030***	-0.001
	(0.001)	(0.002)	(0.004)
Male	0.702***	1.141***	0.543***
	(0.025)	(0.046)	(0.123)
White	0.166***	0.161**	0.344
	(0.043)	(0.082)	(0.224)
Married	-0.139***	-0.237***	-0.404**
	(0.032)	(0.058)	(0.158)
Presence of Teenagers	-0.110***	-0.041	-0.445***
	(0.031)	(0.055)	(0.162)
Income > $100,000	0.371***	0.353***	0.297*
	(0.035)	(0.061)	(0.156)
Income < $25,000	-0.649***	-0.705***	-0.422***
	(0.028)	(0.052)	(0.146)
Observations	39,536		

Notes: Laptop dummy, Firm dummies, and year dummies included
SE in parenthesis. *** significant at 1%, ** at 5%, * at 10%

The parameter on the Intel dummy variable is negative or insignificant in all specifications after other PC/CPU characteristics and CPU advertising amount are taken into account. This result is consistent with the preliminary regression result in Table 4.6 that shows that the purchase probability is actually higher when a PC is powered by AMD chip than by an Intel chip.

Advertising at the CPU level is not valued by consumers in most specifications but has a positive and significant impact in specification (6) after instrumenting for CPU ads. Advertising by a PC firm at the brand level is always positive and significant. In particular, when evaluated at the average value, a one unit increase in PC advertising at the brand level leads to a 0.99 increase in the utility level (specification (5)). When advertising variables are instrumented, there is a larger positive impact of advertising on demand at the PC level: when evaluated at the average value, one unit increase in PC advertising at brand level leads to 1.43 increase in the utility level (specification (6)). Note that a more positive advertising effect after instrumenting for advertising suggests that advertising may be negatively correlated with unobserved product qualities. That is, firms may engage in more advertising when they have lower-quality products. This result appears reasonable given that the low-end processor of Intel Celeron was heavily advertised during the period. The positive impact of PC advertising implies that CPU firms can enhance CPU sales by inducing PC firms to advertise PCs that contain their chip. This provides us with the pro-competitive justifications for Intel's marketing campaign: promoting CPU sales by subsidizing PC advertising. We will test whether the benefit of promoting CPU sales by subsidizing PC advertising exceeds the cost of the marketing subsidy. If not, the marketing campaign may be driven by anticompetitive motives.

Table 4.9: Multinomial Logit Estimates of PC/CPU Demand

Variables		Dependent Variable: ln(Market Share) - ln(Share of Outside Goods) (1)	(2)	(3)	(4)	(5)	(6)
Price		-0.286***	-1.047*	-0.515***	-0.337***	-1.765***	-2.228***
		(0.0855)	(0.609)	(0.0671)	(0.0647)	(0.541)	(0.643)
PC Characteristics	Laptop	-0.178**	0.146			0.410*	0.497*
		(0.0711)	(0.266)			(0.236)	(0.292)
	Older PC	0.906***	0.858***			0.655***	0.582***
		(0.122)	(0.128)			(0.114)	(0.145)
	PC Age	0.126***	0.138***			0.0940***	0.0904***
		(0.0222)	(0.0241)			(0.0215)	(0.0238)
CPU Characteristics	Intel	-0.510***	-0.233	-0.0104	-0.000200	0.502	-2.060
		(0.103)	(0.244)	(0.231)	(0.221)	(0.319)	(1.308)
	ln(Benchmark)	0.371**	0.851**	0.575***	0.415***	1.315***	1.362***
		(0.155)	(0.414)	(0.128)	(0.120)	(0.371)	(0.434)
	ln(Speed)	1.594***	1.731***	1.381***	1.108***	1.625***	1.724***
		(0.106)	(0.149)	(0.0883)	(0.0847)	(0.135)	(0.166)
	Older CPU	-0.0906	-0.0641	0.487***	0.494***	0.237**	0.192
		(0.113)	(0.116)	(0.0922)	(0.0832)	(0.107)	(0.129)
	CPU Age	0.00524	-0.0101	0.0994***	0.0960***	0.0360*	0.0226
		(0.0174)	(0.0206)	(0.0135)	(0.0131)	(0.0186)	(0.0229)
Advertising	CPU ad			-0.0259	0.0945	-0.139	4.186***
				(0.187)	(0.174)	(0.193)	(1.360)
	CPU ad squared			-0.00803	-0.0370	0.0167	-0.868***
				(0.0306)	(0.0278)	(0.0321)	(0.253)
	PC ad			2.352***		2.363***	3.496***
				(0.0830)		(0.0862)	(0.302)
	PC ad squared			-0.294***		-0.304***	-0.581***
				(0.0216)		(0.0223)	(0.133)
Included Controls							
	IV for Price		Included			Included	Included
	IV for Advertising						Included
	PC Advertising				Included		
	PC Advertising Squared				Included		
	PC Firm Fixed Effects	Included	Included	Included	Included	Included	Included
Observations		5,327	5,327	5,327	5,327	5,327	5,327
R-squared		0.498	0.491	0.597	0.649	0.575	0.470
Median (Own) Price Elasticity		-0.336	-1.231	-0.605	-0.396	-2.075	-2.619

Notes:Standard errors are in parenthesis. *** indicates significant at 1%; ** at 5% and * at 10%. All regressions include time dummies.

4.6 Marginal Marketing Revenue

Intel's marketing campaign provided support to PC firms that advertised PCs with Intel chips. One of the benefits of the TAP approach is that it allows us to circumvent modeling Intel's profit maximization problem. This is beneficial both because the test (and model) would become very complicated and the data necessary to estimate such a model do not exist, in particular firm-specific rebate rates are not publicly available. Suppressing time notation, the total marketing/ad revenue (TMR) of Intel from PC sales is

$$TMR = \sum_{c \in \mathcal{J}_{intel}} \sum_{f} (p_c^{CPU} - mc_c)\mathcal{M}s_{fc}(p, a), \tag{4.2}$$

where $\mathcal{J}_{intel}$ is the set of products with an Intel CPU; p_c^{CPU} is the price of CPU c; mc_c is the marginal production cost of CPU c and $s_{fc}(p, a)$ is market share of processor c sold by firm f , which depends on the product price (i.e., PC price, p) and advertising (a).[20]

One issue to address in the PC advertising data is that advertising may involve more than one product. For example, PC firms often engage in general promotions both by platform type and across all platforms (e.g., Acer Laptop Computer; Acer Various Computers) or PC brands may be jointly advertised (e.g., Acer Veriton & Travelmate Computers Combo). We will require a composite measure of advertising expenditures by product that includes all advertising done for that product (so it should include all group advertising). Following Sovinsky Goeree (2008) we compute product advertising expenditures as a weighted average of group advertising for that product where the weights are a function of the number of products in that group. Specifically, suppressing the time subscript, let $\mathcal{G}_j$ be the set of all product groups that include product j with group $\mathcal{H} \in \mathcal{G}_j$. Then composite product ad expenditures for product j are given by

$$a_j^{\text{total pc}} = \sum_{\mathcal{H} \in \mathcal{G}_j} \frac{\lambda_{\mathcal{H}} a_{j\mathcal{H}}^{\text{pc}}}{|\mathcal{H}|}.$$

where the sum is over the different groups that include product j. We also estimate a non-linear specification to allow for increasing or decreasing returns to group advertising.

Intel's marginal revenue from the marketing campaign for firm f is given by

$$MMR_f = \mathcal{M} \sum_{\substack{j,r \in \mathcal{J}_f \cap \\ c \in \mathcal{J}_{intel}}} (p_c^{CPU} - mc_c) \frac{\partial s_j(p, a)}{\partial a_r^{\text{total pc}}}. \tag{4.3}$$

As we discussed previously, we have data on p_c^{CPU} and mc_c. We can use these data together with the demand side estimates to compute the marginal revenue of advertising dollars spent by PC firm f on Intel chips.

[20] The CPU price is the listed price. In practice, many PC firms paid less than listed price as Intel granted discounts on CPU prices to selected firms. As a results our measure of the TMR will be larger than what it would be if we had the purchase price. Thus, our measure of TMR makes the TAP results more stringent than what they would be if we had the purchase price.

4.7 Marginal Marketing Cost

We focus primarily on Intel and its agreements with Dell. This is for two reasons. First, both Intel and Dell were examined separately by antitrust authorities for related antitrust violations. Hence, we have a wealth of information on the amount of Intel's advertising subsidy to Dell, relative to other PC firms, especially during the period 2002 − 2005. Second, antitrust investigations have produced written evidence that Intel's agreements with Dell were intended to exclude their main rival (AMD) from the market, which provides a good test for our model.

We also apply the TAP test to Intel's rebates to HP and Toshiba, who were also investigated by antitrust authorities as part of the Intel case. In addition, we examine the predatory nature of Intel's rebates with Gateway. This latter application serves as a robustness check of our test as there is no evidence that Gateway was involved in anticompetitive behavior with Intel over this period.[21]

4.7.1 Measuring Costs of "Intel Inside" Campaign

AVC of marketing campaign is total marketing subsidy paid by Intel/total advertising by PC firm for PCs with Intel chip. Suppose that PC firms increase advertising expenditures more as Intel increases spending on the marketing program, that is, there are increasing return to scale of the marketing program. Then, average variable cost would be larger than marginal cost and the test may lead to a false positive of predation. In contrast, if PC firms tend to be less responsive to Intel's marketing subsidy as the amount of subsidy increases, average variable cost would be smaller than marginal cost and the test would be lenient.

We use a variety of measures of the average variable cost (AVC) of the marketing/ad campaign as a proxy for the marginal cost of the marketing program, which is in the same spirit as using average variable production cost as a proxy for the marginal cost in Areeda-Turner (1975) test of predatory pricing. The AVC of the marketing/ad campaign is computed by dividing the total

[21] Most large PC firms were involved in the Intel case. Gateway is the largest PC firm that was not under investigation in the Intel case.

dollar amount of the ad subsidy that Intel paid to a PC firm by total PC firm advertising for PCs powered by an Intel chip.

We construct four measures of the observed marginal cost of the marketing/ad campaign based on either actual expenditures paid by Intel (as shown in Table 4.10) or on an assumed percentage rebate rate.

(**MC**$_1$) This measure is constructed as the total payment of Intel to Dell as given in the case files divided by total advertising by Dell for PCs powered by an Intel chip. Table 4.10 Column (4) provides the total payment of Intel's rebates to Dell. Basically Intel provided discounts on CPU prices and sometimes provided a lump-sum payment. At the end of 2001, Intel began a program in which it agreed to give Dell a six percent rebate on all of Dell's CPU purchases (this came to be called the "Meet Competition Program (MCP))."[22] These rebates are treated as a reduction in marketing expenses in accounting by Dell. We can see that rebate rates as well as lump-sum payments have significantly increased between 2002 and 2005. Relative to the operating income, the total amount of rebates were over 70% by 2006 (Column (6)).

(**MC**$_2$) The second measure is the total rebate amount as a percentage of sales of PCs (across all segments sold by Dell) divided by total advertising by Dell for PCs powered by an Intel chip. The rebates to Dell can be used for advertising in all segments including servers. Ours is a model of household demand (due to data limitations) so when we construct MC_2 we assume that the rebates are used in each sector in proportion to sector market shares. Hence MC_2 is MC_1 multiplied by the percentage of Dell's revenue from PC sector in each quarter.

(**MC**$_3$) This measure is a proxy for the marginal cost that we would have in the absence of antitrust documents. The publicly announced subsidy takes the form of a fixed percentage rebate of CPU purchases made by the PC firm with zero lump-sum payments. To compute this measure we assume six percent of all Dell's CPU purchases from Intel are rebated. We recompute the total expenditure of the marketing program to Dell (i.e., we compute column

[22] It was orginally called the "Mother of All Programs (MOAP)."

4 assuming column 1 is always six percent and column 3 is zero). We then divide the total expenditure by total advertising by Dell for PCs powered by an Intel chip.

Table 4.10: Rebate amounts paid by Intel to Dell

		Meet the Competition Program (MCP)				
Quarter	Rebate Rate on CPU purchases (1)	Percentage Rebate Payment (2)	Lump Sum Payment (3)	Total MCP Payments (4)	Dell's Operating Income (5)	MCP % of Operating Income (6)
2002Q1	6%	$61m	-	$61m	$590m	10%
2002Q2	6%	$57m	$3m	$60m	$677m	9%
2002Q3	6%	$59m	$12m	$71m	$758m	9%
2002Q4	6.3%	$77m	$7m	$84m	$819m	10%
2003Q1	6.3%	$91m	$8m	$99m	$811m	12%
2003Q2	6.3%	$106m	$6m	$112m	$840m	13%
2003Q3	6.3%	$105m	$40m	$145m	$912m	16%
2003Q4	7%	$118m	$82m	$200m	$981m	20%
2004Q1	8.7%	$137m	$70m	$207m	$966m	21%
2004Q2	12% + var.	$210m	-	$210m	$1,006m	21%
2004Q3	12% + var.	$250m	-	$250m	$1,095m	23%
2004Q4	12% + var.	$293m	$75m	$368m	$1,187m	31%
2005Q1	12% + var.	$307m	$81m	$388m	$1,174m	33%
2005Q2	12% + var.	$313m	$119m	$432m	$1,173m	37%
2005Q3	12% + var.	$339m	-	$339m	$754m	45%
2005Q4	12% + var.	$423m	$60m	$483m	$1,246m	39%
2006Q1	12% + var.	$405m	$318m	$723m	$949m	76%

Source: p.14 MCP table from "Securities and Exchange Commission vs. Dell", U.S. District Court of Columbia Summary Statement

($\mathbf{MC}_4$) This measure provides a further lower bound on marginal costs. According to the Intel Inside website, Intel would provide a three percent rebate on purchases if marketing featured the Intel logo. To compute this measure we assume the rebate rate is three percent and there are no lump sum payments. Hence MC_4 is half of MC_3.

Notice that the last two measures of marginal cost (MC_3 and MC_4) do not rely on information obtained by antitrust officials but are based only on publicly available information. Therefore, MC_3 and MC_4 can be used to construct the TAP test for other PC manufacturers. In addition, these two measures provide a benchmark as they are computed based on the assumption that the marketing program has been executed as in normal periods or as Intel describes on their website.

The other two measures (MC_1 and MC_2) are computed based on actual payments, which reflect any potential anticompetitive cost increase.

We would like to point out a limitation of applying TAP to Intel. The main issue concerns situations where a predator operates in multiple markets, which makes it difficult to determine how to allocate costs across markets. PC firms are active in markets for home consumers, but also in markets for education, business, and government consumers. Also they sell servers as well as PCs. Fortunately, our advertising data allow us to differentiate PC advertising from advertising for non-PC products. However, we include general promotions at the firm-level as an advertising expenditure in the home segment. If general promotions affect sales in every market segment, then they should be allocated across all segments. In this case, our measure of the average cost of the marketing program is likely to underestimate the actual average cost. The marginal revenue of the marketing program would likewise be underestimated, as we cannot allow for spillovers of advertising across segments (due to data restrictions).

4.8 TAP Results

We apply the TAP test to the case of Intel. Specifically, we consider the potential predatory nature of Intel's marketing program over the period 2002 to 2005. We run TAP based on preliminary demand estimation results from specification (6). Intel's marginal revenue from the marketing campaign for firm f is given by

$$MMR_f = \mathcal{M} \sum_{\substack{j,r \in \mathcal{J}_f \cap \\ c \in \mathcal{J}_{intel}}} (p_c^{CPU} - mc_c) \frac{\partial s_j(p,a)}{\partial a_r^{\text{total pc}}}.$$

where $\mathcal{J}_{intel}$ is the set of Intel CPUs and $\mathcal{J}_f$ are the set of products produced by firm f. We compute the marginal revenue from the advertising campaign earned in that time period (i.e., ours is not a dynamic model). Estimating a dynamic model of demand and computing the associated marginal revenue that arises from a dynamic profit function would introduce considerable difficulties with respect to estimation and data requirements. However, we conducted robustness checks of our TAP results that consider the potential brand building effect of advertising.

The estimated marginal revenue of Intel's marketing subsidy to Dell and its 95 percent confidence interval is given in Table 4.11, along with the marginal cost of the marketing campaign computed in four different ways as described in section 4.2[23]. When the marginal cost measure is above the 95 percent confidence interval for marginal revenue, we conclude that the marketing program is not consistent with profit maximization and, more specifically, that there was an excessive marketing subsidy. Then the test result is "positive". The final column indicates for which of the four measures of marginal cost the TAP result is positive.

Table 4.11: TAP Results of Intel's Marketing Campaign for Dell

	Computed Marginal Revenue					Observed Marginal Cost				TAP Result	# Years	# Years
Time	Est.	95% conf. Interval				(1)	(2)	(3)	(4)	Positive in	(1)	(4)
2002Q1	1.45	(	1.22	1.69	)	34.54	31.12	3.33	1.66	1,2,3	5.9	0.3
2002Q2	0.88	(	0.73	1.02	)	45.62	40.70	3.73	1.87	all	13.0	0.5
2002Q3	1.23	(	1.03	1.43	)	53.31	48.05	5.90	2.95	all	10.9	0.6
2002Q4	2.22	(	1.85	2.58	)	91.86	82.59	13.02	6.51	all	10.4	0.7
										all		
2003Q1	2.95	(	2.46	3.45	)	352.14	315.44	44.78	22.39	all	29.8	1.9
2003Q2	2.38	(	1.98	2.77	)	192.68	171.66	17.50	8.75	all	20.3	0.9
2003Q3	3.02	(	2.52	3.52	)	140.80	126.81	13.61	6.80	all	11.6	0.6
2003Q4	2.92	(	2.45	3.40	)	93.19	82.75	7.57	3.79	all	8.0	0.3
										all		
2004Q1	3.35	(	2.79	3.91	)	682.73	607.55	56.66	28.33	all	51.0	2.1
2004Q2	1.95	(	1.62	2.27	)	345.53	306.09	18.57	9.28	all	44.3	1.2
2004Q3	1.72	(	1.44	2.00	)	133.28	118.52	7.39	3.69	all	19.4	0.5
2004Q4	2.05	(	1.71	2.39	)	192.46	171.46	8.55	4.27	all	23.4	0.5
										all		
2005Q1	6.30	(	5.24	7.35	)	390.02	340.19	15.51	7.75	all	15.5	0.3
2005Q2	4.38	(	3.65	5.11	)	370.16	323.16	11.29	5.65	all	21.1	0.3
2005Q3	4.43	(	3.68	5.17	)	348.36	307.69	15.68	7.84	all	19.7	0.4
2005Q4	5.85	(	4.87	6.83	)	340.69	298.21	13.24	6.62	1,2,3	14.56	0.28

Notes: Unit: \$ (inflation adjusted - base: 2000)

[23] Since the marginal revenue of the marketing program is a linear combination of utility parameters, the confidence interval around the estimate of the marginal revenue can be computed easily. MC_1 is Intel's total payment to Dell divided by the total advertising expenditure of Dell on PCs powered by Intel CPU; MC_2 is MC_1 multiplied by the fraction of revenue from PC sales (= PC sales/(PC sales + server sales); MC_3 is six percent of Intel's revenue from selling CPUs to Dell ($= 0.06 \cdot \mathcal{M} \sum_{c \in \mathcal{J}_{intel}} \sum_{j \in \mathcal{J}_c \cap \mathcal{J}_{Dell}} s_j p_c^{CPU}$); MC_4 is three percent of Intel's revenue from selling CPUs to Dell (so $MC_4 = MC_3/2$).

It is worth to note the following two points. First, the marginal revenue estimates imply that 1$ PC advertising tends to increase Intel's revenue by more than 1$ (except in the second quarter of 2000). Thus, if all of Intel's payment to Dell are used for marketing PCs powered by Intel CPU, then the results suggest that Intel gains more than its cost by subsidizing Dell. However, our measures of marginal cost are much larger than 1, that is, PC advertising expenditure falls short of Intel's marketing subsidy to Dell in reality.[24] This implies that Intel needs to incur more than 1$ to induce Dell to spend 1$ more on advertising. Second, MC_3 and MC_4 are significantly smaller than MC_1 and MC_2. Considering that MC_1 and MC_2 are computed based on the actual payment whereas MC_3 and MC_4 are computed by assuming the rebate rate applied in the normal periods, our results suggest that Intel has paid much more to Dell than described in their Intel Inside marketing campaign.

The TAP results show that marginal cost, for all measures, is above the 95 percent confidence interval of marginal revenue estimates in most periods. Exceptions are when MC_4 is compared to marginal revenue in the beginning of the anticompetitive period and at the end of the anti-competitive period. Even in these cases, MC_4 is larger than the estimated marginal revenue. Comparisons of marginal revenue to the actual-payment-based marginal cost is consistent with predatory marketing.

We do not have the data on Intel's actual payment to other firms and thus cannot compute MC_1 and MC_2 for firms other than Dell. Although, we can compute MC_3 and MC_4 if we assume fixed rebate rates. The anticompetitive charges were mainly about Intel's payment to Dell, but other PC firms such as HP and Toshiba were also involved in the case. Table 4.12 presents the TAP results for HP and Toshiba, respectively. Both the marginal revenue and marginal cost measures tend to be lower in the cases of HP and Toshiba than in the case of Dell. Our TAP

[24] There are a few reasons why marginal costs may be larger than one. First, while, in principle, the marketing subsidy should have been used only for marketing, case files indicate that stock analysts had doubt about the use of the subsidy. The subsidy was actually a huge amount that accounts for a significant portion of operating profits. Second, marketing may be broadly defined to include more than advertising. For example, it may encompass training employees in how to sell PCs.

results show evidence of predation with Toshiba in 2003 and 2004. During this period the TAP results are positive for both measures of marginal cost.

Table 4.12: TAP Results for HP and Toshiba

PC firm	HP						Toshiba					
	Computed MR			Observed MC		TAP Result	Computed MR			Observed MC		TAP Result
Time	Est.	95% conf. Interval		(3)	(4)	Positive in	Est.	95% conf. Interval		(3)	(4)	Positive in
2002Q1	-0.02	(-0.32	0.28)	0.11	0.05		0.13	(0.11	0.15)	0.14	0.07	
2002Q2	1.04	(0.89	1.19)	0.26	0.13		0.26	(0.22	0.29)	0.09	0.04	
2002Q3	1.20	(1.01	1.39)	0.27	0.13		0.66	(0.56	0.76)	0.16	0.08	
2002Q4	-5.68	(-9.93	-1.43)	0.12	0.06	3,4	0.62	(0.53	0.72)	0.45	0.22	
2003Q1	-0.42	(-1.30	0.46)	0.09	0.05		0.34	(0.29	0.40)	1.37	0.68	3,4
2003Q2	0.15	(-0.34	0.64)	0.11	0.05		0.21	(0.18	0.25)	0.82	0.41	3,4
2003Q3	0.36	(-0.25	0.97)	0.17	0.09		0.31	(0.26	0.36)	1.47	0.74	3,4
2003Q4	0.59	(-0.45	1.64)	0.30	0.15		0.44	(0.37	0.51)	1.45	0.72	3,4
2004Q1	-0.07	(-0.80	0.66)	0.14	0.07		0.34	(0.28	0.40)	0.99	0.50	3,4
2004Q2	0.32	(-0.01	0.66)	0.13	0.07		0.36	(0.30	0.42)	0.99	0.50	3,4
2004Q3	0.59	(0.31	0.87)	0.20	0.10		0.43	(0.36	0.49)	0.48	0.24	3
2004Q4	0.37	(-0.04	0.77)	0.17	0.09		0.99	(0.84	1.14)	0.29	0.15	
2005Q1	2.47	(0.80	4.15)	0.17	0.08		1.96	(1.65	2.27)	0.35	0.17	
2005Q2	4.90	(4.12	5.68)	0.33	0.17		1.99	(1.67	2.32)	0.42	0.21	
2005Q3	6.91	(5.90	7.92)	0.57	0.29		2.94	(2.45	3.42)	0.72	0.36	
2005Q4	6.31	(5.18	7.43)	0.41	0.21		2.97	(2.49	3.44)	0.49	0.24	

We are also interested in what TAP would show with other PC firms not involved in the case. Since most PC firms not investigated for the case are small, we consider only Gateway, which has a significant market share but is not involved in the case. Table 4.13 presents the TAP results. We can see that marginal cost measures tend to either fall in the 95 percent confidence interval of marginal revenue or just below the interval. The result provides a nice contrast to the test result for Dell as we do not find excessive advertising for Gateway for both measures of marginal costs in any period. This result also supports the idea that the Intel's marketing program, if conducted as described, is driven by profit maximization rather than anticompetitive purposes for the case of Gateway. Overall, the findings with other PC firms seem to suggest that our model, despite its simplicity, represents demand reasonably well and captures the important features of the marketing program.

Table 4.13: TAP Results for Gateway

PC firm	Gateway					
	Computed Marginal Revenue			Observed Marginal Cost		TAP Result
Time	Est.	95% conf. Interval		(3)	(4)	Positive in
2002Q1	0.56	(0.37	0.76)	0.14	0.07	
2002Q2	0.38	(0.28	0.49)	0.10	0.05	
2002Q3	0.47	(0.38	0.55)	0.16	0.08	
2002Q4	0.78	(0.66	0.90)	0.26	0.13	
2003Q1	0.87	(0.75	1.00)	0.20	0.10	
2003Q2	0.47	(0.40	0.54)	0.20	0.10	
2003Q3	0.50	(0.43	0.58)	0.17	0.08	
2003Q4	0.56	(0.47	0.64)	0.15	0.07	
2004Q1	0.66	(0.56	0.76)	0.78	0.39	3
2004Q2	0.40	(0.34	0.47)	0.50	0.25	3
2004Q3	0.52	(0.44	0.60)	0.37	0.18	
2004Q4	0.15	(0.96	1.34)	2.66	1.33	3
2005Q1	4.52	(3.77	5.28)	6.37	3.18	3
2005Q2	3.59	(3.00	4.19)	2.64	1.32	
2005Q3	4.72	(3.94	5.51)	4.21	2.11	
2005Q4	5.72	(4.80	6.64)	1.79	0.90	

Unit: $ (inflation adjusted - base: 2000)

4.9 Motive, Recoupment, and Efficiency Motives

In this section, we discuss the industry background that speaks to the motives for predation, the prospect of successful predation and recoupment, and dynamic efficiency. Intel is a dominant firm in the CPU industry with about 80 percent of worldwide CPU sales. Its major (and only effective) rival is AMD, holding about 18 percent market share (Mercury Research (2007)). In 1999 and 2003, respectively, AMD introduced two new chips, the Athlon for personal computers and the Opteron for servers. The threat of new, high-performance processors from AMD may have induced Intel to engage in anticompetitive actions. These events provide the motive for Intel's predatory behavior. Indeed, many jurisdictions in the world accused Intel of using various anticompetitive tactics against AMD starting in 2002.

We are particularly interested in Intel's marketing subsidies. Predation involves short-run profit sacrifice and long-run recoupment. The TAP test is used to establish short-run profit sacrifice. However, we now turn to industry characteristics to examine the ease (or difficulty) with which Intel could successfully drive AMD out of the market and recoup lost profits by maintaining market power for a sufficiently long period after AMD's exit.

There are a number of factors that make long-run recoupment of profits likely to be successful in the CPU industry. To remain as a valid competitor in a rapidly changing, high-technology industry like the CPU industry, firms need to secure constant cash flows and keep investing in innovation. The CPU industry is capital-intensive, hence firms will incur substantial costs to construct and maintain manufacturing plants (called "fabs"). If a firm does not have sufficient internal funding, it must obtain external funding at market rates. According to industry experts, Intel is able to fund its fabs with revenue, while AMD must secure funding at market rates, which significantly raises AMD's cost of capital. Furthermore, obtaining external financing is complicated due to agency problems. Typically investors require firms to show a positive prospect of future profits, which is often based on current performance. Predation would make the future prospect of the prey look lower (and potentially negative) and ultimately induce it to exit the market. Thus, predation in the CPU market would be consistent with the long-purse (deep-pocket) theory of predation.

Furthermore, since firms are continuously innovating, they may be uncertain about how consumers will react to new products. New processors can have different characteristics possibly appealing to a different market segment from current customers. As mentioned before, the beginning of the anticompetitive use of the marketing program coincides with AMD's introduction of high-performance chips. By engaging in predatory behavior, Intel could send a (wrong) signal about the demand for new chips, which is consistent with the demand signaling theory (test-market theory) of predation.

Lastly, economies of scale exist in the CPU industry. The substantial investment in plants and technologies are sunk. Therefore, a firm needs to secure a certain amount of sales in order to recover the sunk costs and stay in business. It is easier for a dominant firm to exclude a rival and prevent new entrants in the presence of economies of scale. In this sense, predation is likely to be successful in driving AMD out of a market and Intel is likely to keep high profit margins for a sufficiently long time.

The CPU industry is inviting to predatory behavior for these reasons, and Intel is an incumbent with a dominant market share. Given that Intel's recoupment is very likely as a monopolist due to high entry barriers and that predation can successfully lead to exclude AMD in the CPU industry, showing sacrifice of short term profits would support that the marketing program is predatory.

The TAP test examines if the return on advertising (i.e., how it impacts demand) is high enough to justify marketing expenditures (as these are directed at increasing demand). Short-term profit sacrifice may be justified by dynamic efficiency reasons. Although the cost-based approach is widely used to show profit sacrifice in predatory pricing cases, pricing below cost does not necessarily mean the behavior is predatory. Short-term profit sacrifice can be rationalized by potential dynamic efficiency reasons such as learning-by-doing, promotional purposes (e.g., introductory prices), or network externalities. Our demand model includes only the current, short-term effect of advertising, hence the potential long-run benefit of the marketing program is not taken into account. However, just as the efficiency reasons for pricing below marginal cost are not usually applicable to an already dominant, incumbent firm with a large customer base, here too an unprofitable advertising subsidy by Intel is not easily justified by efficiency reasons. Intel should already have achieved an efficient scale of operation, so learning-by-doing does not seem to justify short-term profit sacrifice. Given that Intel has been present for a long time and consumers already know about Intel and that the anticompetitive actions have been going on for four years, promotional motives are an unlikely explanation for short-term profit sacrifice. In addition, network externalities are not strong in the CPU market. For example, PC purchase guides, such as *Consumer Reports,* do not list the size of the customer base using Intel processors as an important factor for consumers to consider when purchasing a PC.

Our main concern is the brand-loyalty-building effect of the marketing program. Advertising is generally believed to build goodwill and this may be a reason for Intel to invest in marketing at the expense of short-term profits. Notice that, this incentive is constant across all periods, while the predatory motive is more pronounced during the period of AMD's new chip introductions. To consider this the TAP test includes two measures of marginal cost (MC_3 and MC_4) that serve as competitive benchmark as they are based on listed rebate rates that would have been applied prior to 2001/2002. In contrast, MC_1 and MC_2 are based on actual payment post AMD's introduction of the new chips. Hence, these measures would include brand loyalty building incentives plus anticompetitive motives, while MC_3 and MC_4 would be driven only by brand loyalty building incentive. We find that MC_1 and MC_2 (based on actual payment) are much larger than MC_3 and MC_4 for Dell. The results suggest that an anticompetitive motive induced Intel to sacrifice even more short-term profits as the difference between marginal marketing revenue and marginal costs are much larger when using MC_1 and MC_2. Also it is worth to note that, if advertising can establish strong brand loyalty, predatory marketing can be even more harmful as it would work as an endogenous entry barrier, deterring further entries and making recoupment even more likely.

4.10 Conclusion

Price and quantity are not the only strategic variables that can be used for anticompetitive purposes. Advertising is another important strategic variable commonly employed by firms. However, antitrust authorities typically try to establish anticompetitiveness through pricing, but do not address the strategic use of advertising and, more generally, marketing campaigns. While the heart of the anticompetitive actions of Intel was their Intel-Inside marketing program, considerations of advertising/marketing predation where not at the forefront of the antitrust case. In this paper we focus on non-price anticompetitive behavior arising from marketing/advertising with a focus on the Intel case.

Our paper proposes a "Test of Advertising Predation" (TAP) that can be used to detect non-price predatory behavior. We provide a general test based on the presumption that, if a firm's marketing campaign is not predatory, marketing expenses should be profit maximizing and so should result in sufficient increased product demand to justify costs. To construct TAP, first we model consumer's demand for PCs from which we infer demand for CPU processors. Specifically, we estimate a random-coefficient model of demand for a PC-CPU, where the coefficients on PC and CPU characteristics and advertising vary with demographics[25]. Second, we compute Intel's marginal revenue from the marketing subsidy using the demand side estimates. That is, we compute the marginal revenue of advertising dollars spent on Intel chips at the firm or product level[26]. The marginal revenue of the marketing program depends on the parameters of consumer utility (including advertising), CPU price and marginal manufacturing cost.

Test results suggest short-term profit sacrifice by Intel, supporting the predatory use of the Intel Inside campaign. To rationalize the short-term profit sacrifice, there should be something Intel can gain from the marketing program other than increasing CPU sales by boosting the willingness-to-pay for a PC. Antitrust authorities found evidence that the marketing subsidy is paid on anticompetitive condition that limits the use of AMD processors. This condition aimed at driving AMD out of a market and the prospect of future profit as a monopolist as a result would have rationalized the short-term profit sacrifice.

Our method can be used to guide antitrust authorities in future cases as it provides a general framework for testing for anticompetitive use of marketing campaigns. Computing the test requires little extra estimation over the typical demand estimation usually undertaken by antitrust authorities. Furthermore, the advertising data necessary to estimate the model parameters is usually not so difficult to obtain. Thus, TAP is practical and easy to implement. Applying to the Intel case, this paper shows that our test can be used to show the predatory use of advertising/marketing. In addition, the benefit of looking at the advertising side is that, unlike predatory pricing which accompanies low price in the short term, predatory advertising/marketing does not

[25] In previous literature that estimates CPU demand, it is generally assumed that final consumers directly purchase the CPU. We think it is more realistic to model consumers' choice of a CPU-PC combination. In addition, since we are interested in the effect of PC advertising on CPU demand, and PC advertising does not directly affect CPU demand, we model a consumer's discrete choice over CPU-PC combinations.

[26] We do not model strategic decisions of PC firms and Intel. This makes the test we develop more stringent. Rather, PC firms' CPU choices are assumed to simply reflect consumers' demand, and not affected by the marketing campaign. Intel was accused of giving refunds to PC firms in the Intel Inside marketing program on the exclusionary condition that they limit the use of AMD chips. This implies the marketing program would affect PC firms' CPU choice and, hence, its anticompetitiveness would be even larger.

have a clear benefit for consumers, even in the short term. In the long run, predatory marketing can be harmful if it has a long-lasting effect by establishing goodwill, which may become an endogenous entry barrier for potential competitors.

References

1. Ackerberg, D. (2003), "Advertising, learning, and consumer choice in experience good markets: an empirical examination", *International Economic Review*, Vol. 44(3), 1007-1040.

2. Areeda, P. and D.F. Turner (1975) "Predatory Pricing and Related Issues Under Section 2 of the Sherman Act", *Harvard Law Review*, 88, 697-733.

3. Berry, S., J. Levinsohn and A. Pakes. (1995) "Automobile prices in market equilibrium," *Econometrica* 63(4), 841-890.

4. Berry, S., J. Levinsohn and A. Pakes. (1998) "Differentiated products demand systems from a combination of micro and macro data: The new car market", NBER Working Paper, w6481.

5. Besanko, D., U. Doraszelski, and K.Yaroslav (2010) "The Economics of Predation: What Drives Pricing When There is Learning-by-Doing?," Unpublished Manuscript.

6. Bolton, P., J. Broadley, and M. Riordan (2000) "Predatory Policy: Strategic Theory and Legal Policy ", *Georgetown Law Review*, 88, 2239-2330.

7. Bolton, P., J. Broadley, and M. Riordan (2001) "Predatory Pricing: Response to Critique and further Elaboration" , *Georgetown Law Journal,* 89, 2496-2529.

8. Bureau of Labor Statistics, All Urban Consumers CPI-U, U.S. city average.

9. Chamberlain, G. (1987) "Asymptotic efficiency in estimation with conditional moment restrictions", *Journal of Econometrics*, 34(3), 305-334.

10. Chen, Y. and W. Tan (2007) "Predatory Advertising: Theory and Evidence in the Pharmaceutical Industry," Working Paper.

11. Eizenberg A. (2011) "Upstream Innovation and Product Variety in the United States Home PC Market, " Working Paper.

12. Ellison, G. and S. F. Ellison (2011) "Strategic Entry Deterrence and the Behaviour of Pharmaceutical Incumbents Prior to Patent Expiration," *American Economic Journal: Microeconomics*, 3(1), 1-36.

13. Genesove, D. and W. P. Mullin (2006) "Predation and Its Rate of Return: The Sugar Industry, 1887-1914, " *Rand Journal of Economics.*

14. Goettler, R. and B. R. Gordon (2009) "Competition and innovation in the microprocessor industry: Does AMD spur Intel to innovate more? "Working paper, Columbia University, New York.

15. Goldberg, P. (1995) "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry"*Econometrica*, 63(4), 891-951.

16. Gowrisankaran, G. and M. Rysman (2007) "Dynamics of consumer demand for new durable consumer goods, " Working paper, Boston University, Boston.

17. Gordon, B. (2009). "A Dynamic Model of Consumer Replacement Cycles in the PC Processor Industry, " Unpublished Working Paper, Columbia Business School.

18. Granitz, E. and B. Klein (1996) "Monopolization by 'Raising Rivals' Costs': the Standard Oil Case " . *Journal of Law and Economics* 39, 1-47.

19. In-Stat (2005) "Intel Manufacturing Capacity and Die Cost ".

20. In-Stat (2005) "Intel Rosetta Stone: Intel Processor Shipments, Forecasts, Technology, and Roadmaps ".

21. Moon, Y. E. and C. Darwall (2005) "Inside Intel Inside, " *Harvard Business School Case* 502-583.

22. Nevo A. (2001) “Measuring Market Power in the Ready-to-Eat Cereal Industry”, *Econometrica*, Vol. 69(2), 307-342.

23. Prince, J. T. (2008) “Repeat Purchase Amid Rapid Quality Improvement: Structural Estimation of Demand for Personal Computers,” *Journal of Economics and Management Strategy* , 17(1), 1-33.

24. Reynaert M. and F. Verboven (2014) “Improving the Performance of Random Coefficients Demand Models: the Role of Optimal Instruments”, *Journal of Econometrics*, 179(1), 83-98.

25. Salgado, H. (2008a) “Dynamic Firms Conduct and Market Power: The Computer Processors Industry under Learning-by-Doing,” Working Paper, University of California at Berkeley.

26. Salgado, H. (2008b) “Brand Loyalty, Advertising and Demand for Personal Computer Processors: The Intel Inside Effect,” Working Paper, University of California at Berkeley.

27. Slade, M. E. (2005) “Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition,” Working Paper.

28. Snider, C. (2009) ”Predatory Incentives and Predation Policy: The American Airlines Case”, Working Paper.

29. Song, M. (2007) “Measuring Consumer Welfare in the CPU Market: An Application of the Pure Characteristics Demand Model,” *RAND Journal of Economics*, 38, 429-446.

30. Sovinsky Goeree, M. (2008) “Limited Information and Advertising in the US Personal Computer Industry,” *Econometrica*, 1017-1074.

31. Weiman, D. and R. Levin (1994) “Preying for monopoly? The case of Southern Bell Telephone Company, 1894-1912” , *Journal of Political Economy* 102, 103-26.

Appendix 4.A

Table 4.A.1: Product Cross-Reference from Processor Core to Brand Name (i.e. Marketing Name) in Sample (Q1:2002 - Q4:2005)

Platform		Processor Core	Brand Name	Speed (Frequency: MHz)
Desktop	Mainstream	Willamette	Pentium 4	1300 - 2000
		Northwood		1600 - 3400
		Prescott		2260 - 3800
		Smithfield*	Pentium D	2667 - 3200
	Value	Tualatin	Pentium III	1000 - 1400
			Celeron	900 - 1400
		Willamette	Celeron	1500 - 2000
		Northwood		1600 - 2800
		Prescott	Celeron D	2133 - 3460
Mobile	Mainstream	Northwood	Mobile Pentium 4-M	1200 - 2600
		Prescott	Mobile Pentium 4	2300 - 3460
		Banias	Pentium M	1200 - 1800
		Dothan		1300 - 2267
	Value	Tualatin	Mobile Celeron	1000 - 1330
			Mobile Pentium III-M	866 - 1333
		Northwood	Mobile Celeron	1400 - 2500
		Banias	Celeron M	1200 - 1500
		Dothan		1200 - 1700
	Low-Power	Tualatin LV	Mobile Pentium III-M	733 - 1000
		Tualatin ULV		700 - 933
		Tualatin LV	Mobile Celeron	650 - 1000
		Tualatin ULV		650 - 800
		Banias LV	Pentium M	1100 - 1300
		Banias ULV		900 - 1100
		Dothan LV		1400 - 1600
		Dothan ULV		1000 - 1300
		Banias ULV	Celeron M	600 - 900
		Dothan ULV		900 - 1000

Notes: * Dual-core processor
Low-power mobile PCs are mini-notebook, tablet, and ultraportables.
(LV: low-voltage; ULV: ultra-low-voltage)

Chapter 5

Curriculum Vitae

Personal Details

Name:	András Imre Péchy
Date of Birth:	19th November 1984
Place of Birth:	Budapest, Hungary
Nationality:	Hungarian

Education

2011 September - 2016 July	PhD Studies, Zurich Graduate School of Economics
	University of Zurich, Switzerland
	Supervisor: Prof. Dr. Michelle Sovinsky
2007 September - 2009 July	Master of Science in Economics
	HEC School, University of Lausanne, Switzerland
2004 September - 2007 July	Bachelor of Science in Economics
	HEC School, University of Lausanne, Switzerland